Uniwersytet Warszawski
Wydział Biologii

Jakub Czarnecki
Nr albumu: 262418

Analiza struktury i funkcji chromidów wybranych gatunków bakterii z rodzaju *Paracoccus* (*Alphaproteobacteria*)

Rozprawa doktorska
w zakresie nauk biologicznych
dyscyplinie biologii

Praca wykonana pod kierunkiem
prof. dr. hab. Dariusza Bartosika
oraz dr. hab. Łukasza Dziewita
Zakład Genetyki Bakterii, Instytut Mikrobiologii,
Wydział Biologii, Uniwersytet Warszawski

Warszawa, październik 2015

Oświadczenie kierującego pracą

Oświadczam, że niniejsza praca została przygotowana pod moim kierunkiem i stwierdzam, że spełnia ona warunki do przedstawienia jej w postępowaniu o nadanie stopnia doktora nauk biologicznych w zakresie biologii.

Data Podpis kierującego pracą

Oświadczenie autora pracy

Świadom odpowiedzialności prawnej oświadczam, że niniejsza rozprawa doktorska została napisana przeze mnie samodzielnie i nie zawiera treści uzyskanych w sposób niezgodny z obowiązującymi przepisami.

Oświadczam również, że przedstawiona praca nie była wcześniej przedmiotem procedur związanych z uzyskaniem stopnia doktora w innej jednostce.

Oświadczam ponadto, że niniejsza wersja pracy jest identyczna z załączoną wersją elektroniczną.

Data Podpis autora pracy

Słowa kluczowe

chromid, plazmid, *Paracoccus* spp., metylotrofia

Tytuł pracy w języku angielskim

Analysis of structure and functions of chromids of selected *Paracoccus* spp. strains

Źródła finansowania

1. Grant NCN Opus 2013/09/B/NZ1/00133 “**Analiza struktury i funkcji chromidów występujących w bakteriach z rodzaju *Paracoccus* (*Alphaproteobacteria*)**”. Okres realizacji: 2014-2017.

2. Grant MNiSW Iuventus Plus 2011 011471 – „**Porównawcza genomika strukturalna i funkcjonalna *Paracoccus aminovorans* JCM 7685 i *Paracoccus aminophilus* JCM 7686 – metylotroficznych bakterii zdolnych do rozkładu związków toksycznych**”. Okres realizacji: 2012-2014.

3. Grant MNiSW Iuventus 008670 „**Identyfikacja, analiza i wykorzystanie modułów genetycznych warunkujących rozkład związków toksycznych do konstrukcji szczepów bakteryjnych użytecznych w procesach bioremediacji**”. Okres realizacji: 2010-2011.

4. Grant Uniwersytetu Warszawskiego DSM 140000/501/86-107409 “**Molekularna charakterystyka szlaków metylotroficznych bakterii *Paracoccus aminovorans***”. Okres realizacji: 2014-2015.

5. Grant Uniwersytetu Warszawskiego BW100043 „**Analiza ładunku genetycznego plazmidu pKON1 Paracoccus kondratievae**”. Okres realizacji: 2011-2012.

Podziękowania

Serdecznie dziękuję mojemu promotorowi, **prof. dr. hab. Dariuszowi Bartosikowi**, za bezcenną pomoc na wszystkich etapach przygotowania mojej rozprawy doktorskiej, troskę o mój rozwój naukowy i osobisty, optymizm oraz nieustające wsparcie moich, kosztownych nieraz, pomysłów.

Dr. hab. Łukaszowi Dziewitowi za możliwość wzięcia udziału w prowadzonych przez niego projektach badawczych, impulsy prowadzące do mojego zainteresowania metabolizmem bakterii oraz chłodne spojrzenie na problemy Wszechświata.

Dr Jadwidze Baj za przyjaźń, za to, że zawsze znajdowała czas, aby mnie wysłuchać i zachwycić się rzeczami, których nikt inny nie rozumie, za krytyczne spojrzenie na moje teksty pisane, a przede wszystkim za to, że dzięki Niej zostałem mikrobiologiem.

Dr Karolinie Skorek za to, że nie ma potrzeby pisać, za co jej dziękuję.

Wszystkim moim przyjaciołom z Instytutu Mikrobiologii UW, a przede wszystkim **dr Magdalenie Szuplewskiej, mgr. Robertowi Laskowi, dr Adriannie Raczkowskiej, mgr Marii Puzynie, mgr Marcie Nieckarz, mgr. Łukaszowi Kowalskiemu, Katarzynie Kuźmicz, Aleksandrze Brozio, mgr Anecie Lis i mgr. Michałowi Szymczakowi** za przyjemność współistnienia i współpracy.

Spis treści

Lista prac włączonych do rozprawy doktorskiej

Wyniki niniejszej rozprawy zostały opublikowane w następujących pracach oryginalnych artykułach naukowych:

1. **Jakub Czarnecki**, Łukasz Dziewit, Łukasz Kowalski, Magdalena Ochnio, Dariusz Bartosik. 2015. Maintenance and genetic load of plasmid pKON1 of *Paracoccus kondratievae*, containing a highly efficient toxin-antitoxin module of the *hipAB* family. Plasmid 80:45-53.

2. Łukasz Dziewit, **Jakub Czarnecki**, Daniel Wibberg, Monika Radlińska, Paulina Mrozek, Michał Szymczak, Andreas Schlüter, Alfred Pühler, Dariusz Bartosik. 2014. Architecture and functions of a multipartite genome of the methylotrophic bacterium *Paracoccus aminophilus* JCM 7686, containing primary and secondary chromids. BMC Genomics 15:124.

3. Łukasz Dziewit*, **Jakub Czarnecki***, Emilia Prochwicz, Daniel Wibberg, Andreas Schlüter, Alfred Pühler, Dariusz Bartosik. 2015. Genome-guided insight into the methylotrophy of *Paracoccus aminophilus* JCM 7686. Frontiers in Microbiology 6:852.

* dwóch równorzędnych pierwszych autorów publikacji

Streszczenie

Podstawowa informacja genetyczna, determinująca funkcje niezbędne do wzrostu i podziałów komórkowych bakterii, zawarta jest w chromosomie. Cechą charakterystyczną genomów wielu bakterii jest występowanie replikonów pozachromosomowych. Replikony pozachromosomowe, zaliczane najczęściej do grupy plazmidów, niosą informację dodatkową, która nie jest niezbędna do prawidłowego funkcjonowania komórek gospodarzy. U niektórych bakterii doszło jednak do przeniesienia, w wyniku rekombinacji, części genów metabolizmu podstawowego z chromosomu do współwystępujących plazmidów, co doprowadziło do powstania niezbędnych komórce replikonów (chromidów) o właściwościach typowych zarówno dla plazmidów, jak i chromosomu. Chociaż chromidy występują dość powszechnie (m.in. u wielu bakterii ważnych z punktu widzenia człowieka), zjawisko podziału niezbędnej informacji genetycznej między kilka replikonów oraz wpływ tych rearanżacji na biologię i ewolucję bakterii, nie zostały dotąd dokładnie zbadane. Brak spójnej, jednoznacznej definicji „niezbędności" genu oraz niedoskonałość metod umożliwiających identyfikację replikonów niezbędnych, bardzo często stwarzają trudności w zaklasyfikowaniu replikonów pozachromosomowych do grupy plazmidów bądź chromidów.

Bakterie należące do rodzaju *Paracoccus* (*Alphaproteobacteria*), zasiedlające powszechnie rozmaite środowiska, mają genomy o strukturze wieloreplikonowej. W niniejszej pracy podjęto próbę identyfikacji chromidów w dwóch szczepach *Paracoccus* (*P. kondratievae* NCIMB 13773 i *P. aminophilus* JCM 7686) oraz zdefiniowania informacji genetycznej determinującej właściwości tych replikonów.

Do identyfikacji chromidów wykorzystano zjawisko niezgodności systemów replikacyjnych (ang. *incompatibility*), co pozwoliło na wskazanie replikonów, których nie można usunąć z komórki. Badania przeprowadzone w *P. kondratievae* NCIMB 13773 wykazały jednak, że kryterium nieusuwalności nie może być jedyną podstawą wyróżniania chromidów, ze względu na częste występowanie w replikonach pozachromosomowych efektywnych systemów stabilizujących, powodujących posegregacyjną eliminację komórek pozbawionych danego replikonu. W trakcie badań wykazano, że wysoka stabilność replikonu pKON1 *P. kondratievae* (95 049 pz) wynika z obecności właśnie tego typu systemu, kodującego toksynę (HipA) i antytoksynę (HipB). Ponieważ replikon ten nie zawiera genów metabolizmu podstawowego, został on zaliczony do grupy plazmidów.

Stosując analogiczne podejście metodyczne, w genomie *P. aminophilus* JCM 7686 zidentyfikowano dwa replikony pozachromosomowe, ważne dla wzrostu i prawidłowego funkcjonowania komórek tej bakterii. Pierwszy z nich (pAMI5, 294 013 pz) to replikon nieusuwalny w testowanych warunkach. Ponieważ nie można go usunąć, nawet po wyeliminowaniu wpływu obecnego w jego obrębie systemu toksyna-antytoksyna, jest prawdopodobne, że niesie on geny metabolizmu podstawowego, bez których bakteria nie może żyć. Drugi replikon (pAMI6, 206 582 pz) udało się usunąć z komórek gospodarza, jednak uzyskany w ten sposób szczep charakteryzował się znacznym spowolnieniem wzrostu na podłożu pełnym oraz utratą zdolności do wzrostu na podłożu minimalnym. Zatem pAMI6 niesie geny istotne dla wzrostu, co niewątpliwie wpływa na konkurencyjność *P. aminophilus* w środowisku naturalnym.

Obydwa replikony niosą liczne geny występujące w chromosomach blisko spokrewnionych bakterii. Zaproponowano, aby zarówno pAMI5, jak i pAMI6, zaliczyć do grupy chromidów. Jednak, ze względu na zaobserwowane różnice między nimi, pAMI5 nazwano chromidem pierwszorzędowym (obligatoryjnie niezbędnym), a pAMI6 – chromidem drugorzędowym (fakultatywnie niezbędnym).

W trakcie badań wykazano również, że informacja genetyczna zawarta w obrębie chromidów *P. aminophilus* JCM 7685 determinuje charakterystyczne cechy metaboliczne tego szczepu. Obydwa replikony niosą pulę genów, których produkty są zaangażowanych w metylotrofię (zdolność do wykorzystania związków C1 - zredukowanych związków węgla niezawierających wiązań węgiel-węgiel), która umożliwia zasiedlenie i przeżycie komórek gospodarza w jego naturalnej niszy, tj. glebie zanieczyszczonej *N,N*-dimetyloformamidem (związek C1). Ponieważ analizowane chromidy zawierają informację genetyczną umożliwiającą zajmowanie typowej dla danej bakterii niszy ekologicznej, stanowią one tzw. replikony determinujące styl życia (ang. *lifestyle-determining replicons*).

Summary

The presence of extrachromosomal replicons is a characteristic feature of many bacterial genomes. While the genetic information essential for growth and division of bacterial cells is located within the chromosome, the extrachromosomal replicons, usually classified as plasmids, can provide functions that are critical for the survival of a bacterium in a specific environment, however, they are not indispensable for the viability of the host cells.

Comparative genomic studies revealed that in many bacterial genomes some housekeeping genes have been transferred from the chromosome into the co-occurring plasmids. This phenomenon has led to the generation of essential extrachromosomal replicons, called chromids, sharing features of both chromosomes and plasmids. Although chromids have been identified in the genomes of many well-studied pathogenic and biotechnologically-relevant strains, the biological significance of such genome rearrangement has not yet been thoroughly investigated. The unclear concept of “essentiality” of some genetic information and the lack of a unified methodology for the identification of essential replicons make it difficult to formulate an exact definition of a chromid.

The representatives of the genus *Paracoccus* (*Alphaproteobacteria*) are widespread both in pristine and polluted environments. The genetic information of many *Paracoccus* spp. strains is dispersed between chromosome and numerous extrachromosomal replicons. The main aim of this study was to identify essential extrachromosomal replicons (chromids) within two *Paracoccus* strains (*P. kondratievae* NCIMB 13773 and *P. aminophilus* JCM 7686) and to define the genetic information responsible for the specific features of these replicons.

Chromids have been distinguished from plasmids by the attempts to remove individual extrachromosomal replicons from a cell with the use of the target-oriented, replicon curing technique. However, such an analysis performed in *P. kondratievae* NCIMB 13773 revealed that such a method of determination of replicon “essentiality” is not sufficient to distinguish chromids from plasmids. The reason is that many replicons, including pKON1 (95 049 bp) of *P. kondratievae*, carry very efficient toxin-antitoxin systems, which cause post-segregational elimination of cells deprived of the replicon from bacterial population. Such replicons, which do not carry any essential genes, should be classified as plasmids.

In the genome of *P. aminophilus* JCM 7686, two replicons (out of eight extrachromosomal replicons of this strain) were found to be crucial for the viability or proper growth of the strain. One of them, pAMI5 (294 013 bp), could not be removed from the host cells by incompatibility in any of the tested conditions, even after the elimination of the influence of the pAMI5-encoded toxin-antitoxin system . This result suggests that pAMI5 carries housekeeping genes and its loss is lethal to the cell. The second replicon, pAMI6 (206 582 bp), could be easily removed from the JCM 7686 cells. However, the obtained pAMI6-less strain showed slower growth rate in rich media and was completely unable to grow in minimal media. Thus, pAMI6 encodes genes essential for normal growth of *P. aminophilus*, and its loss undoubtedly affects the competitiveness of the strain in its natural environment.

pAMI5 and pAMI6 carry numerous genes conserved in the chromosomes of many closely related bacterial species. It was proposed to classify both these replicons as chromids. Based on the differences in their significance for the host cells, pAMI5 was defined as a primary chromid (obligatorily essential), and pAMI6 as a secondary chromid (facultatively essential).

Further analyses revealed that the *P. aminophilus* JCM 7686 chromids determine metabolic capabilities characteristic for this strain, and thus they can be considered as lifestyle-determining replicons. Both chromids carry genes involved in methylotrophy (metabolic capability that allows the use of C1 compounds, e.g. methane, methanol, methylamine, as the sole source of carbon and energy), which is the key feature allowing *P. aminophilus* to inhabit its natural niche, i.e. soil polluted with *N,N*-dimethylformamide (C1 compound).

1. Wstęp

U większości bakterii, całkowita informacja genetyczna, istotna dla przebiegu podstawowych procesów biologicznych, skumulowana jest w pojedynczym replikonie, określanym jako **chromosom**. Oprócz chromosomu, bakterie niosą często dodatkowe replikony – **plazmidy**, o bardzo zróżnicowanej strukturze i właściwościach. Niektóre z nich determinują jedynie funkcje związane z własną replikacją, transferem i stabilnym utrzymywaniem w komórce, dlatego też są postrzegane przez wielu badaczy jako „pasożyty molekularne". Inne, ze względu na obecność genów o charakterze adaptacyjnym, odgrywają istotną rolę, umożliwiają bowiem lepsze dostosowanie gospodarzy do zmiennych warunków środowiska. Replikony te można jednak usunąć z komórek bakteryjnych, co nie wpływa negatywnie na żywotność ich gospodarzy.

Genomy wielu bakterii (np. *Vibrio cholerae*, *Brucella* spp., *Agrobacterium tumefaciens*, *Sinorhizobium meliloti*) mają bardziej złożoną strukturę, bowiem występują w nich dodatkowe duże replikony, które, podobnie jak chromosomy, zawierają geny metabolizmu podstawowego. W odróżnieniu od chromosomu**,** dodatkowe niezbędne replikony niosą stosunkowo niewielką ilość niezbędnej informacji genetycznej. Z tego powodu niektóre z nich nazwano **chromosomami drugorzędowymi**, stosując jednocześnie nazwę **chromosom pierwszorzędowy** dla określenia głównego replikonu w komórce bakteryjnej. Inne dodatkowe niezbędne replikony, z uwagi na ich podobieństwo do plazmidów, zaliczono do **megaplazmidów**, podkreślając tym samym znaczne rozmiary tych replikonów.

Tak duża rozbieżność w definiowaniu pozachromosomowych replikonów niezbędnych, skłoniła badaczy do poszukiwania jednolitych kryteriów, umożliwiających pełniejsze ich scharakteryzowanie. Jedna z ciekawszych propozycji zakładała wyróżnienie nowej grupy replikonów - **chromidów**, równocennej chromosomom i plazmidom, do której włączono zarówno „chromosomy drugorzędowe", jak i niezbędne „megaplazmidy" [Harrison i wsp., 2010].

Szczegółowe analizy z zakresu genomiki porównawczej umożliwiły wyciągnięcie ogólnych wniosków na temat właściwości chromidów oraz powszechności ich występowania. Okazało się, że te niezbędne replikony pozachromosomowe występują u około 10% bakterii, których genomy zostały zsekwencjonowane, a do ich powstania doszło, niezależnie, w różnych grupach filogenetycznych. Są to pochodne plazmidów, dlatego zawierają typowe dla tych replikonów systemy replikacyjne. Niosą one jednocześnie informację genetyczną pochodzenia chromosomowego (w tym, niezbędne do funkcjonowania komórek gospodarza, geny metabolizmu podstawowego), która uległa międzyreplikonowemu transferowi, w wyniku procesów rekombinacyjnych.

Sekwencje nukleotydowe chromosomów i chromidów mają podobne właściwości (średnia zawartość par GC oraz częstość wykorzystywania kodonów w sekwencjach kodujących), co wskazuje na ich długą wspólną ewolucję w danym gospodarzu [Harrison i wsp., 2010; Poirion, 2014].

Przyjęcie koncepcji chromidu zasadniczo zmienia obraz struktury genomów bakteryjnych. W jej świetle, prokarioty te mają tylko jeden chromosom, z silnie konserwowanym systemem replikacyjnym, zależnym od funkcji białka inicjacyjnego DnaA

[Mackiewicz i wsp., 2004]. Replikony definiowane jako chromidy, ze względu na ich „plazmidowe korzenie", powinny być zaliczane, wraz z plazmidami, do grupy **replikonów pozachromosomowych**.

Postawienie ostrej granicy pomiędzy „chromidami" a „plazmidami" jest jednak problematyczne. Replikony pozachromosomowe trudno podzielić jednoznacznie na „zbędne" oraz „niezbędne", bowiem wiele z nich wykazuje pośredni wpływ na tempo wzrostu i *fitness* gospodarza [Dziewit i wsp., 2014; Stasiak i wsp., 2014]. Ponadto, zaobserwowano, że niektóre konserwowane replikony, niemające wpływu na tempo wzrostu w warunkach laboratoryjnych, mogą być niezbędne w naturalnej niszy ekologicznej [np.: Finan i wsp., 2001; Chain i wsp., 2006; Harrison i wsp., 2010; Petersen i wsp., 2013; Frank i wsp., 2015a; Frank i wsp, 2015b; Soora i wsp., 2015]. Nie zawierają one genów metabolizmu podstawowego pochodzenia chromosomowego, zatem nie mogą być zidentyfikowane jedynie na podstawie trzech „sztywnych" kryteriów bioinformatycznych zaproponowanych przez Harrisona i wsp. [2010]. Stąd pojawiły się propozycje rozszerzenia definicji chromidów (chromidy niezbędne w każdych warunkach i chromidy fakultatywnie niezbędne) oraz włączenia analiz eksperymentalnych, jako kluczowego etapu w procedurze ich identyfikacji [Petersen i wsp., 2013; Dziewit i wsp., 2014, Dziewit i Bartosik, 2015].

Dogodnym obiektem do tego typu analiz są bakterie z klasy *Alphaproteobacteria*, które powszechnie zawierają w swoich genomach duże replikony pozachromosomowe. Do tej klasy należą m.in. *Rhodobacter sphaeroides* 2.4.1, pierwsza bakteria, u której stwierdzono genom zawierający więcej niż jeden replikon niezbędny [Suwanto i Kaplan, 1989], oraz gatunki z rodzajów *Agrobacterium*, *Rhizobium* i *Sinorhizobium*, których liczne megareplikony scharakteryzowano na poziomie molekularnym.

Do *Alphaproteobacteria* należy również rodzaj *Paracoccus*, którego przedstawiciele bytują w różnorodnych środowiskach – od naturalnych środowisk glebowych i morskich, po środowiska skażone, oczyszczalnie ścieków i bioreaktory, a nawet tkanki człowieka [np.: Urakami i wsp., 1990; Siller i wsp., 1996; Lipski i wsp., 1998; Tsubokura i wsp., 1999; Funke i wsp., 2004; Lee i wsp., 2004; Liu i wsp., 2006]. Bakterie te charakteryzuje duża plastyczność metabolizmu, przejawiająca się m.in. zdolnością wielu szczepów do wzrostu zarówno heterotroficznego, z wykorzystaniem licznych związków organicznych jako źródła węgla i energii, jak i chemolitoautotroficznego, z udziałem zredukowanych związków siarki, wodoru cząsteczkowego czy jonów żelaza (II) jako źródeł energii [Kelly i wsp., 2006]. Bakterie z tego rodzaju znane są również z możliwości wykorzystania różnych końcowych akceptorów elektronów; wiele z nich to warunkowe tlenowce, zdolne do oddychania azotanowego. Szczepy *Paracoccus* spp. mające zdolność rozkładu ksenobiotyków znalazły praktyczne zastosowanie w procesach bioremediacji [np.: Sun i wsp., 2013].

Do niedawna, w bazie danych GenBank (NCBI) dostępna była całkowita sekwencja nukleotydowa genomu tylko jednego szczepu z rodzaju *Paracoccus* – *P. denitrificans* Pd1222. Genom tej bakterii obejmuje trzy replikony: chromosom 1 (2,85 Mpz), chromosom 2 (1,73 Mpz) oraz plazmid 1 (0,65 Mpz), z których dwa ostatnie, wg kryteriów Harrisona i wsp., powinny być zaliczone do grupy chromidów [Harrison i wsp., 2010]). Obecnie dostępna jest również sekwencja genomu *P. aminophilus* JCM 7686, którą ustalono w Zakładzie Genetyki Bakterii Uniwersytetu Warszawskiego (ZGB UW). Genom tej bakterii, składający się z chromosomu (3,61 Mpz) oraz aż ośmiu replikonów pozachromosomowych (o wielkości w zakresie od 5,6 do 438 kpz), jako doskonały model do analiz chromidów, stał się jednym z obiektów badań prowadzonych w ramach przedstawionej rozprawy doktorskiej.

Badania te zainicjowały kompleksowe analizy genomiczne kilkunastu innych gatunków *Paracoccus*, które prowadzone są aktualnie w ZGB UW. Ich dotychczasowych efektem jest ustalenie kompletnej sekwencji genomu *P. aminovorans* JCM 7685, składającego się z chromosomu (3,11 Mpz) i trzech replikonów pozachromosomowych (740 kpz, 185 kpz i 4,2 kpz), oraz sekwencji kilkudziesięciu replikonów pozachromosmowych pochodzących z innych szczepów *Paracoccus* spp. Jednym z nich jest replikon pKON1 (95 kpz) z *P. kondratievae* NCIMB 13773, którego analiza stanowi jeden z wątków niniejszej rozprawy.

2. Cel pracy

Celem badań stanowiących podstawę niniejszej rozprawy doktorskiej była identyfikacja niezbędnych replikonów pozachromosomowych w genomach *P. aminophilus* JCM 7686 i *P. kondratievae* NCIMB 13773, oraz wyróżnienie i scharakteryzowanie informacji genetycznej, determinującej specyficzne właściwości poszczególnych replikonów.

3. Omówienie uzyskanych wyników

Przeprowadzone badania koncentrowały się na dwóch głównych wątkach tematycznych. Pierwszy obejmował klasyfikację i charakterystykę poszczególnych replikonów pozachromosomowych *P. aminophilus* JCM 7686 oraz replikonu pKON1 *P. kondratievae* NCIMB 13773, który, jak sugerowały wyniki badań wstępnych, wykazywał pewne właściwości chromidu, natomiast drugi dotyczył analizy funkcji zidentyfikowanych niezbędnych replikonów (chromidów). W przedstawionym niżej opisie odnoszono się do nieopublikowanych dotąd danych na temat replikonów pozachromosomowych *P. aminovorans* i *P. denitrificans*, co pozwoliło na przedstawienie wyników rozprawy w znacznie szerszym kontekście genomicznym.

3.1. Identyfikacja chromidów w bakteriach z rodzaju *Paracoccus*

> **Publikacja 1:** *Maintenance and genetic load of plasmid pKON1 of* Paracoccus kondratievae*, containing a highly efficient toxin-antitoxin module of the* hipAB *family.*
>
> **Publikacja 2:** *Architecture and functions of a multipartite genome of the methylotrophic bacterium* Paracoccus aminophilus *JCM 7686, containing primary and secondary chromids.*

Jak wcześniej wspomniano, chromidy, jako replikony niosące geny metabolizmu podstawowego, są niezbędne do przeżycia swoich gospodarzy [Harrison i wsp., 2010]. Podjęto więc próby usunięcia poszczególnych replikonów pozachromosomowych z komórek analizowanych szczepów *Paracoccus* spp. Wykorzystano do tego celu zjawisko niezgodności (ang. *incompatibility*), określane jako niezdolność występowania w jednej komórce dwóch replikonów niosących identyczne bądź blisko spokrewnione systemy replikacyjne. Do przeprowadzenia analiz konieczne było skonstruowanie serii mobilizowalnych plazmidów wahadłowych (*E. coli-Paracoccus* spp.), zawierających systemy replikacyjne poszczególnych replikonów pozachromosomowych *P. aminophilus* oraz replikonu pKON1 *P. kondratievae*. Uzyskane plazmidy wahadłowe wprowadzano następnie, drogą koniugacji, do komórek *P. aminophilus* bądź *P. kondratievae*, w celu usunięcia ich naturalnych replikonów.

Genom *P. kondratievae* nie został w pełni poznany, dlatego analizom poddano tylko jeden replikon tej bakterii - pKON1 (95 049 pz). Replikonu tego nie udało się usunąć z komórek w wyniku niezgodności, istniało więc przypuszczenie, że ma on naturę chromidu. W genomie pKON1 zidentyfikowano segment DNA o dużej identyczności sekwencji z fragmentem chromosomu 1 *P. denitrificans* Pd1222, a także inne geny odnajdowane w chromosomach wielu bakterii. Szczegółowe analizy eksperymentalne wykazały jednak, że replikon ten nie zawiera genów metabolizmu podstawowego, a jego dużą stabilność determinuje bardzo wydajny system toksyna-antytoksyna (TA) z rodziny *hipAB*. Zaobserwowano również, że nawet blisko spokrewnione moduły TA działają z różną

wydajnością, i nie zawsze powodują posegregacyjną eliminację wszystkich komórek pozbawionych niosącego je replikonu.

Obserwacje poczynione podczas analizy pKON1 pokazują, że identyfikacja chromidów nie może polegać wyłącznie na próbach ich usunięcia drogą niezgodności. Bardzo ważne jest również uwzględnienie i wyeliminowanie wpływu systemów stabilizujących, np. typu toksyna-antytoksyna lub restrykcji i modyfikacji, których obecność nadaje pozory „niezbędności" danego replikonu. Podkreśla to wagę i znaczenie analiz eksperymentalnych, jakim należy poddawać poszczególne replikony pozachromosomowe [Czarnecki i wsp., 2015].

Dalsze analizy modułu *hipAB* pKON1 wykazały, że system ten wydajnie stabilizuje heterologiczne wektory w różnych szczepach bakterii z klasy *Alphaproteobacteria* (brak detekcji komórek bezplazmidowych). Znalazł on więc zastosowanie, jako kaseta genetyczna, do konstrukcji wektorów stabilnie utrzymywanych w komórkach bez konieczności stosowania presji selekcyjnej antybiotyku [wyniki nieopublikowane, uzyskane przez autora niniejszej rozprawy].

Podczas badań nad plazmidem pKON1 zwrócono także uwagę na nietypową strukturę genomu tego replikonu, który aż w ok. 30% składa się z kompletnych bądź defektywnych elementów transpozycyjnych oraz pseudogenów powstałych w wyniku transpozycji. Co ciekawe, geny, których ciągłość została przerwana podczas transpozycji, nie mają swojej kontynuacji w obrębie pKON1. Obserwacje te sugerują, że nie tylko transpozycja, lecz również inne zdarzenia rekombinacyjne (w tym rekombinacja homologiczna pomiędzy kopiami TE występującymi w różnych replikonach genomu *P. kondratievae*) odcisnęły swoje piętno na strukturze tego replikonu.

Większość badań prezentowanych w niniejszej rozprawie przeprowadzono z wykorzystaniem szczepu *P. aminophilus* JCM 7686. W pierwszej kolejności, uzyskano kompletną sekwencję genomu tej bakterii. Jak wcześniej wspomniano, był to drugi szczep z rodzaju *Paracoccus*, dla którego ustalono pełną mapę fizyczną genomu, składającego się z chromosomu oraz ośmiu replikonów pozachromosomowych. Analizując funkcje poszczególnych replikonów wykazano, że dwa z nich mają szczególne znaczenie dla żywotności bądź wzrostu komórek gospodarza. Były to: (i) pAMI5 (294 013 pz), replikon obligatoryjnie niezbędny (nazwany chromidem pierwszorzędowym), którego nie można było usunąć w żadnych warunkach, nawet po wyeliminowaniu wpływu obecnego w jego obrębie systemu TA, oraz (ii) pAMI6 (206 582 pz), replikon fakultatywnie niezbędny (nazwany chromidem drugorzędowym), którego usunięcie powodowało upośledzenie wzrostu komórek *P. aminophilus* na podłożu pełnym LB oraz utratę zdolności do wzrostu na podłożach minimalnych. Chromid drugorzędowy niesie więc informację genetyczną warunkującą konkurencyjność komórek gospodarza w środowisku naturalnym [Dziewit i wsp., 2014].

Stosując analogiczne podejście eksperymentalne, dokonano również identyfikacji chromidów w genomach dwóch innych szczepów *Paracoccus* spp – *P. denitrificans* Pd1222 i *P. aminovorans* JCM 7685 [dane nieopublikowane zgromadzone przez autora niniejszej rozprawy]. Okazało się, że w obu przypadkach występują tam nieusuwalne chromidy pierwszorzędowe (chromosom 2 *P. denitrificans* – 1,73 Mpz i replikon pAMV3 *P. aminovorans* – 740 kpz), których systemy replikacyjne należą do tej samej grupy niezgodności, a ich genomy wykazują pewien stopień konserwacji sekwencji. Co ciekawe, replikony te mają inną strukturę niż chromid pierwszorzędowy pAMI5 *P. aminophilus*. Obecność tak różnych chromidów w szczepach z tego samego rodzaju jest unikatowa i wskazuje na dużą plastyczność i zmienność genomów *Paracoccus* spp. (wg analiz przeprowadzonych przez Harrisona i wsp. chromidy są zwykle konserwowane w obrębie

całych rodzajów taksonomicznych bakterii [Harrison i wsp., 2010]). W genomie *P. aminovorans* JCM 7685 zidentyfikowano także chromid drugorzędowy pAMV1 (185 kpz), który niesie około 40% genów wspólnych z chromidem drugorzędowym pAMI6 z *P. aminophilus* JCM 7685. Chromidy drugorzędowe nie występują natomiast w genomie *P. denitrificans* Pd1222 [dane nieopublikowane zgromadzone przez autora niniejszej rozprawy].

3.2. Analiza funkcji zidentyfikowanych chromidów

> **Publikacja 2:** *Architecture and functions of a multipartite genome of the methylotrophic bacterium* Paracoccus aminophilus *JCM 7686, containing primary and secondary chromids.*
>
> **Publikacja 3:** *Genome-guided insight into the methylotrophy of Paracoccus aminophilus JCM 7686.*

Informację genetyczną chromidu pierwszo- i drugorzędowego *P. aminophilus* poddano szczegółowym analizom porównawczym. W pierwszej kolejności poszukiwano genów pochodzenia chromosomowego, które determinują specyficzne właściwości tych replikonów. Zaobserwowano, że pAMI5 i pAMI6 zawierają różnej wielkości segmenty DNA, które są konserwowane w chromosomach blisko spokrewnionych gatunków bakterii. W genomie pAMI5 występują m.in. geny odpowiedzialne za syntezę NAD *de novo* i gen *alkB* zaangażowany w naprawę DNA [Dziewit i wsp., 2014]. Aby jednak sprawdzić, czy odpowiadają one za niezbędność pAMI5, konieczne jest przeprowadzenie szczegółowych analiz eksperymentalnych. Trzeba także pamiętać, że w niektórych przypadkach, zidentyfikowane eksperymentalnie niezbędne geny kodują hipotetyczne białka o nieznanej funkcji [np. diCenzo i wsp., 2013].

Analizując pAMI6, zwrócono uwagę na obecność genu, kodującego białko z grupy transporterów jonów amonowych, a więc związków, które stanowią jedyne źródło azotu w wykorzystanym w badaniach podłożu minimalnym [Dziewit i wsp., 2014]. Kontynuowane przeze mnie badania w tym zakresie wykazały jednak, że za utratę zdolności wzrostu na tego typu podłożach odpowiadają dwa inne geny pAMI6 (usytuowane u innych bakterii w chromosomach), kodujące podjednostki reduktazy siarczanowej asymilacyjnej (*cysND*). Enzym ten pozwala na asymilację siarki z zawartych w podłożu siarczanów. Wprowadzenie genów *cysND* do szczepu pozbawionego pAMI6 przywraca zdolność do wzrostu na podłożu minimalnym, jednak kinetyka wzrostu sugeruje, że analizowany chromid niesie równie inne geny ważne dla wzrostu *P. aminophilus* [dane nieopublikowane, uzyskane przez autora niniejszej rozprawy].

Chromidy, obok genów metabolizmu podstawowego, zawierają również informację genetyczną o charakterze adaptacyjnym, umożliwiającą przebieg charakterystycznych dla danego szczepu procesów metabolicznych. Analizy pAMI5 i pAMI6 wykazały, że obydwa replikony warunkują przystosowanie *P. aminophilus* JCM 7686 do zasiedlanej przez niego niszy ekologicznej. Bakteria ta, podobnie jak *P. aminovorans* JCM 7685, została wyizolowana z gleby zanieczyszczonej *N,N*-dimetyloformamidem (DMF), który może stanowić dla tego szczepu jedyne źródło węgla i energii [Urukami i wsp., 1990]. *P. aminophilus* jest zdolny do wzrostu z wykorzystaniem również innych związków C1 (zredukowanych związków węgla niezawierających wiązań węgiel-węgiel), na przykład

trimetyloaminy i metanolu, zaliczany jest więc do grupy mikroorganizmów metylotroficznych. Zarówno pAMI5, jak i pAMI6 okazały się niezbędne dla metabolizmu metylotroficznego. pAMI6 zawiera geny umożliwiające rozkład metylowanych amin (trimetyloaminy, dimetyloaminy oraz metyloaminy), a pAMI5 niesie gen jednego z kluczowych enzymów, niezbędnych do przeprowadzania cyklu serynowego, który umożliwia asymilację węgla ze związków C1. Należy podkreślić, że nie tylko chromidy, lecz także inne replikony pozachromosomowe *P. aminophilus* są zaangażowane w metylotrofię. Na przykład niewielki plazmid pAMI2 zawiera geny kodujące podjednostki *N*,*N*-dimetyloformamidazy (DMFazy), enzymu umożliwiającego utlenianie DMF [Dziewit i wsp., 2010]. Informacja genetyczna pozostałych replikonów pozachromosomowych *P. aminophilus* również może mieć wpływ na konkurencyjność tego szczepu w szczególnych warunkach środowiska. Plazmidy pAMI1, pAMI4 i pAMI8 prawdopodobnie pełnią istotną rolę w transporcie i metabolizmie wielu związków węgla, a pAMI8 zawiera również geny *vir*, które, być może, odpowiadają za niezidentyfikowane dotąd interakcje *P. aminophilus* z komórkami eukariotycznymi [Dziewit i wsp., 2014].

Inne poznane do tej pory chromidy pierwszorzędowe *Paracoccus* spp. mają odmienny ładunek genetyczny od pAMI5. Szczególnie ciekawym przykładem jest chromosom 2 (wg obecnej nomenklatury – chromid) *P. denitrificans*, który koduje białko DnaA inicjujące replikację chromosomu 1 oraz podjednostki polimerazy DNA. W tym układzie replikacja chromosomu 1 jest więc w pełni uzależniona od obecności chromidu. W genomie pAMV3 *P. aminovorans* zidentyfikowano inne geny pochodzenia chromosomowego, m.in. gen czynnika translacji EF-P oraz geny podjednostek reduktazy difosforanów rybonukleozydów, biorącej udział w syntezie dNTP [dane nieopublikowane zgromadzone przez autora niniejszej rozprawy]. Chromid drugorzędowy pAMV1 *P. aminovorans*, podobnie jak pAMI6 *P. aminophilus*, warunkuje metylotrofię. Co ciekawe, replikon ten zawiera nie tylko geny zaangażowane w utlenianie metylowanych amin, lecz również komplet geny enzymów cyklu serynowego, które w *P. aminophilus* zlokalizowane są w obrębie chromosomu [dane nieopublikowane, uzyskane przez autora niniejszej rozprawy; Dziewit i wsp., 2015]. Chromid pAMV1, mimo podobieństw sekwencji z pAMI6, nie zawiera genów *cysDN*, a fenotyp obserwowany po jego utracie (zahamowanie wzrostu na podłożach minimalnych) wynika z obecności w jego obrębie genów zaangażowanych w wykorzystanie licznych źródeł węgla, stosowanych w podłożach minimalnych (L-arabinozy, bursztynianiu, fumaranu, asparaginianu) [dane nieopublikowane, uzyskane przez autora niniejszej rozprawy].

4. Podsumowanie uzyskanych wyników

Do najważniejszych wyników uzyskanych w trakcie realizacji badań wchodzących w skład niniejszej rozprawy doktorskiej zaliczyć należy:

(i) uzyskanie mapy fizycznej złożonego genomu *P. aminophilus* JCM 7686, zawierającego osiem replikonów pozachromosomowych

(ii) identyfikację dwóch replikonów pozachromosomowych (pAMI5 i pAMI6) niosących informację genetyczną istotną z punktu widzenia podstawowych procesów życiowych *P. aminophilus* oraz ich zaklasyfikowanie do grup chromidów pierwszo- i drugorzędowych

(iii) identyfikację w obrębie chromidów potencjalnych genów metabolizmu podstawowego oraz genów adaptacyjnych, zaangażowanych w procesy metaboliczne istotne dla konkurencyjności *P. aminophilus* w naturalnej niszy ekologicznej (geny zaangażowane w metylotrofię).

(iv) identyfikację w genomie *P. aminophilus* JCM 7686 genów determinujących metylotrofię i ustalenie przebiegu szlaków metylotroficznych

(v) wykrycie w genomie pKON1 efektywnego systemu stabilizującego typu toksyna-antytoksyna z rodziny *hipAB*, którego obecność uniemożliwia usunięcie plazmidu z komórek *P. kondratievae*

(vi) zaobserwowanie znaczących różnic w funkcjonowaniu pokrewnych modułów *hipAB*, sugerujących różne cele komórkowe homologicznych toksyn

(vii) ustalenie nietypowej struktury plazmidu pKON1, ukształtowanej w wyniku licznych zdarzeń rekombinacyjnych.

Przedstawione w rozprawie badania stanowią pierwszą próbę identyfikacji i charakterystyki chromidów występujących w bakteriach z rodzaju *Paracoccus*. Jak wspomniano, stanowią one wycinek większego projektu genomicznego, którego realizacja pozwoli wyciągnąć bardziej ogólne wnioski na temat zmienności i ewolucji chromidów tych bakterii. Zdefiniowane w tej pracy chromidy *P. aminophilus* będą stanowiły odniesienie do szczegółowych analiz z zakresu genomiki porównawczej.

5. Spis literatury

diCenzo G, Milunovic B, Cheng J, Finan TM. 2013. The $tRNA^{arg}$ gene and *engA* are essential genes on the 1.7-Mb pSymB megaplasmid of *Sinorhizobium meliloti* and were translocated together from the chromosome in an ancestral strain. J Bacteriol 195:202-12.

Chain PS, Denef VJ, Konstantinidis KT, Vergez LM, Agulló L, Reyes VL, Hauser L, Córdova M, Gómez L, González M, Land M, Lao V, Larimer F, LiPuma JJ, Mahenthiralingam E, Malfatti SA, Marx CJ, Parnell JJ, Ramette A, Richardson P, Seeger M, Smith D, Spilker T, Sul WJ, Tsoi TV, Ulrich LE, Zhulin IB, Tiedje JM. 2006. *Burkholderia xenovorans* LB400 harbors a multi-replicon, 9.73-Mbp genome shaped for versatility. Proc Natl Acad Sci USA 103:15280-7.

Czarnecki J, Dziewit L, Kowalski L, Ochnio M, Bartosik D. 2015. Maintenance and genetic load of plasmid pKON1 of *Paracoccus kondratievae*, containing a highly efficient toxin-antitoxin module of the hipAB family. Plasmid 80:45-53.

Dziewit L, Czarnecki J, Wibberg D, Radlinska M, Mrozek P, Szymczak M, Schlüter A, Pühler A, Bartosik D. 2014. Architecture and functions of a multipartite genome of the methylotrophic bacterium *Paracoccus aminophilus* JCM 7686, containing primary and secondary chromids. BMC Genomics 15:124.

Dziewit L, Czarnecki J, Prochwicz E, Wibberg D, Schlüter A, Pühler A, Bartosik D. 2015. Genome-guided insight into the methylotrophy of *Paracoccus aminophilus* JCM 7686. Front Microbiol 6:852.

Finan TM, Weidner S, Wong K, Buhrmester J, Chain P, Vorhölter FJ, Hernandez-Lucas I, Becker A, Cowie A, Gouzy J, Golding B, Pühler A. 2001. The complete sequence of the 1,683-kb pSymB megaplasmid from the N_2-fixing endosymbiont *Sinorhizobium meliloti*. Proc Natl Acad Sci U S A 98:9889-94.

Frank O, Göker M, Pradella S, Petersen J. 2015a. Ocean's twelve: Flagellar and biofilm chromids in the multipartite genome of *Marinovum algicola* DG898 exemplify functional compartmentalization. Environ Microbiol doi: 10.1111/1462-2920.12947.

Frank O, Michael V, Päuker O, Boedeker C, Jogler C, Rohde M, Petersen J. 2015b. Plasmid curing and the loss of grip – the 65-kb replicon of *Phaeobacter inhibens* DSM 17395 is required for biofilm formation, motility and the colonization of marine algae. Syst Appl Microbiol 38:120-7.

Funke G, Frodl R, Sommer H. 2004. First comprehensively documented case of *Paracoccus yeei* infection in a human. J Clin Microbiol 42:3366-8.

Harrison PW, Lower RP, Kim NK, Young JP. 2010. Introducing the bacterial 'chromid': not a chromosome, not a plasmid. Trends Microbiol 18:141-8

Kelly DP, Rainey FA, Wood, AP. 2006. The genus *Paracoccus*, w: Prokaryotes, ed. M Dworkin, S Falkow, E Rosenberg, KH Schleifer, E Stackebrandt (New York, NY; Springer), 232–249.

Lee JH, Kim YS, Choi TJ, Lee WJ, Kim YT. 2004. *Paracoccus haeundaensis* sp. nov., a Gram-negative, halophilic, astaxanthin-producing bacterium. Int J Syst Evol Microbiol 54:1699-702.

Lipski A, Reichert K, Reuter B, Sproer C, Altendorf K. 1998. Identification of bacterial isolates from biofilters as *Paracoccus alkenifer* sp. nov. and *Paracoccus solventivorans* with emended description of *Paracoccus solventivorans*. Int. J. Syst. Bacteriol. 48:529–536.

Liu XY, Wang BJ, Jiang CY, Liu SJ. 2006. *Paracoccus sulfuroxidans* sp. nov., a sulfur oxidizer from activated sludge. Int J Syst Evol Microbiol 56:2693-5.

Mackiewicz P, Zakrzewska-Czerwinska J, Zawilak A, Dudek MR, Cebrat S. 2014. Where does bacterial replication start? Rules for predicting the *oriC* region. Nucleic Acids Res 32:3781-91.

Petersen J, Frank O, Göker M, Pradella S. 2013. Extrachromosomal, extraordinary and essential – the plasmids of the *Roseobacter* clade. Appl Microbiol Biotechnol 97:2805-15.

Poirion O. 2014. Discrimination analytique des génomes bactériens. Rozprawa doktorska. Ecole Centrale de Lyon.

Siller H, Rainey FA, Stackebrandt E, Winter J. 1996. Isolation and characterization of a new gram-negative, acetone-degrading, nitrate-reducing bacterium from soil, *Paracoccus solventivorans* sp. nov. Int. J. Syst. Bacteriol. 46:1125–1130.

Soora M, Tomasch J, Wang H, Michael V, Petersen J, Engelen B, Wagner-Döbler I, Cypionka H. 2015. Oxidative stress and starvation in *Dinoroseobacter shibae*: the role of extrachromosomal elements. Front Microbiol 6:233.

Stasiak G, Mazur A, Wielbo J, Marczak M, Zebracki K, Koper P, Skorupska A. 2014. Functional relationships between plasmids and their significance for metabolism and symbiotic performance of *Rhizobium leguminosarum* bv. trifolii. J Appl Genet 55:515-27.

Sun M, Luo Y, Teng Y, Christie P, Jia Z, Li Z. 2013. Tenax TA extraction to understand the rate-limiting factors in methyl-β-cyclodextrin-enhanced bioremediation of PAH-contaminated soil. Biodegradation 24:365-75.

Suwanto A, Kaplan S. 1989. Physical and genetic mapping of the *Rhodobacter sphaeroides* 2.4.1 genome: presence of two unique circular chromosomes. J Bacteriol 171:5850-9.

Tsubokura A, Yoneda H, Mizuta H. 1999. *Paracoccus carotinifaciens* sp. nov., a new aerobic gram-negative astaxanthin-producing bacterium. Int. J. Syst. Bacteriol. 49:277–282.

Urakami T, Araki H, Oyanagi H, Suzuki K, Komagata K. 1990. *Paracoccus aminophilus* sp. nov. and *Paracoccus aminovorans* sp. nov., which utilize *N,N*-dimethylformamide. Int. J. Syst. Bacteriol. 40:287–291.

Załączniki

Publikacja 1

Jakub Czarnecki, Łukasz Dziewit, Łukasz Kowalski, Magdalena Ochnio, Dariusz Bartosik

Maintenance and genetic load of plasmid pKON1 of *Paracoccus kondratievae*, containing a highly efficient toxin-antitoxin module of the *hipAB* family

Plasmid, 2015, 80:45-53

Impact Factor: 1,578
Punktacja MNiSW: 20

Plasmid 80 (2015) 45–53

Contents lists available at ScienceDirect

Plasmid

journal homepage: www.elsevier.com/locate/yplas

Maintenance and genetic load of plasmid pKON1 of *Paracoccus kondratievae*, containing a highly efficient toxin–antitoxin module of the *hipAB* family

Jakub Czarnecki, Lukasz Dziewit, Lukasz Kowalski, Magdalena Ochnio, Dariusz Bartosik *

Department of Bacterial Genetics, Institute of Microbiology, Faculty of Biology, University of Warsaw, Miecznikowa 1, 02-096 Warsaw, Poland

ARTICLE INFO

Article history:
Available online 6 March 2015
Communicated by Eva Top

Keywords:
Toxin–antitoxin
hipAB
Plasmid genomics
Insertion sequence
Paracoccus kondratievae
Alphaproteobacteria

ABSTRACT

Paracoccus kondratievae NCIMB 13773^T, isolated from the maize rhizosphere, carries a large (95,049 bp) plasmid pKON1, whose structure has been significantly influenced by transposition. Almost 30% of the plasmid genome is composed of complete or truncated insertion sequences (ISs), representing seven IS families. The ISs are accompanied by numerous genes and gene clusters commonly found in bacterial chromosomes, encoding, among others, (i) a putative type III secretion system of the *Rhizobiales*-T3SS family, (ii) a type I restriction–modification system associated with the anti-codon nuclease (ACNase) gene *prrC* and (iii) OstA and OstB proteins involved in trehalose synthesis. The backbone of pKON1 is composed of replication and partitioning modules conserved in several large alphaproteobacterial replicons, including secondary chromid pAMI6 of *Paracoccus aminophilus* JCM 7686 and chromosome 2 (chromid) of *Rhodobacter sphaeroides* 2.4.1. pKON1 also contains a toxin–antitoxin system of the *hipAB* family, whose presence precludes removal of the plasmid from bacterial cells. This system, unlike two other related *hipAB*-family *loci* originating from plasmid pAMI8 and the chromosome of *Paracoccus aminophilus* JCM 7686, is highly efficient and permits very stable maintenance of a heterologous replicon in various hosts.

1. Introduction

Many bacterial genomes have multipartite structures. They often contain numerous extrachromosomal replicons, including (i) plasmids, the most common type of autonomous replicon, which play a crucial role in horizontal gene transfer and (ii) chromids, a newly outlined group of indispensable replicons, sharing features of both plasmids and chromosomes (Harrison et al., 2010). These replicons are stably maintained in the bacterial population due to the presence of various stabilization systems, acting at different stages of cell cycle.

Multireplicon genome architecture is especially common among *Alphaproteobacteria*, including members of the genus *Paracoccus*. Good examples of bacteria with such a genome structure are *Paracoccus aminophilus* JCM 7686, which, besides a single chromosome, carries two chromids and six plasmids (three in the size range 118–438 kb), and carotenoid-producing strains of *Paracoccus marcusii*, *Paracoccus aestuarii* and *Paracoccus haeundaensis*, carrying numerous low-molecular-weight plasmids (Dziewit et al., 2014; Maj et al., 2013).

In silico DNA sequence analysis has revealed that toxin–antitoxin (TA) systems are the most abundant type of stabilization system among extrachromosomal replicons of *Paracoccus* spp. Within the *P. aminophilus* genome, there are 11 toxin–antitoxin pairs representing the families *relBE/parDE*

* Corresponding author. University of Warsaw, Faculty of Biology, Institute of Microbiology, Department of Bacterial Genetics, Miecznikowa 1, Warsaw, 02-096, Poland. Fax: (+48)225541404.
E-mail address: bartosik@biol.uw.edu.pl (D. Bartosik).

http://dx.doi.org/10.1016/j.plasmid.2015.02.003
0147-619X/

(5 TA systems), *phd-doc* (1), *ccdAB* (1), *hipAB* (2), and 2 hybrid systems (*phd-vapC*) (Dziewit et al., 2014). Plasmids of the aforementioned carotenoid-producing *Paracoccus* strains contain, in total, seven TA modules representing five different TA families (Maj et al., 2013).

TA *loci* have a very simple structure, generally encoding only two elements: (i) a stable toxin protein, which binds to a specific cellular target, and (ii) an unstable antitoxin (protein or antisense RNA), which counteracts the toxin (Hayes, 2003). The functionality of TA systems relies on the different stabilities of the two components. Once a bacterium loses the TA-containing replicon, the labile antitoxin is degraded, and the released toxin may interact with a cellular target. This results in growth arrest or death of the replicon-less segregant cells, which increases the number of replicon-containing cells in the bacterial population (Hayes, 2003).

TA systems are of special interest because they are highly abundant in bacterial chromosomes and may also perform various physiological functions in, for example, the induction of persister cell formation, the general stress response and the regulation of biofilm formation (Unterholzner et al., 2013). It is highly probable that extrachromosomally-located TA cassettes were originally chromosomal *loci* that have been adapted as stabilizing systems following their transfer to mobile genetic elements.

The TA systems have been classified in five types, depending on the nature and mode of action of the antitoxins. Type II, IV and V systems encode proteic antitoxins, while type I and III systems rely on RNA antitoxins (Goeders and Van Melderen, 2014). TA systems of all five types are encoded by bacterial chromosomes, while bacterial plasmids carry exclusively type I and type II systems. Type II systems are the most abundant and they produce diverse toxins, classified in 12 super-families (Leplae et al., 2011).

Due to their specific properties, TA systems have been utilized for the construction of versatile tools for basic research and biotechnology (Stieber et al., 2008). They are commonly used in the creation of positive-selection cloning vectors and protein expression vectors, which can be stably maintained in a bacterial population without the use of antibiotics (Unterholzner et al., 2013).

The majority of analyzed TA systems improve the stability of different replicons, but they are not completely effective, which precludes their biotechnological use (*e.g.* Pecota et al., 1997).

In this study, we found that plasmid pKON1 of *Paracoccus kondratievae* NCIMB 13773^T carries a highly efficient TA system, which may be used for the construction of stable shuttle vectors for *Paracoccus* spp. – a metabolically diverse group of bacteria that includes many strains of both scientific and biotechnological value.

2. Materials and methods

2.1. Bacterial strains, plasmids and culture conditions

The following bacterial strains were used in this study: *Paracoccus kondratievae* NCIMB 13773^T (Doronina and Trotsenko, 2000), *P. kondratievae* NCIBM 131773R (rifampicin resistant derivative of the NCIMB 13773^T strain) (Dziewit et al., 2012), *P. kondratievae* UW300 (pKON1-less derivative of NCIBM 131773R) (this study), *Paracoccus aminophilus* JCM 7686R (Bartosik et al., 2002a), *P. aminophilus* UW200 (pAMI8-less derivative of JCM 7686R) (Dziewit et al., 2014), *Paracoccus aminovorans* JCM 7685R (Bartosik et al., 2002a), *Paracoccus homiensis* DSM 17862R (Dziewit et al., 2012), *Paracoccus pantotrophus* KL100 (Bartosik et al., 2002a) and *Paracoccus versutus* UW225 (Bartosik et al., 1993). All strains used in experiments were rifampicin-resistant derivatives of the wild-type strains. The *Escherichia coli* strains TG1 (Gibson, 1984) and DH5α (Hanahan, 1983) were employed for plasmid construction. All the strains, except one, were grown in Luria–Bertani (LB) medium at 30 °C (*Paracoccus* spp.) or 37 °C (*E. coli*). The exception, *P. homiensis*, was grown in marine broth (Difco) at 30 °C. Where necessary, growth media were supplemented with antibiotics at the following concentrations: kanamycin, 50 μg ml^{-1}; rifampicin, 50 μg ml^{-1}; ampicillin, 100 μg ml^{-1}; tetracycline, 1 (*Paracoccus* spp.) or 20 μg ml^{-1} (*E. coli*). The *Paracoccus* spp. strains formed colonies on solid media after 48 h of incubation. The plasmids used and constructed in this study are described in Table 1.

Table 1
Plasmids used and constructed in this study.

Plasmid	Characteristics	Reference
pKON1	Natural plasmid of *P. kondratievae* NCIMB 13773^T	This study
pABW2	Tc^r, *ori* pMB1, mobilizable cloning vector, *lacZα*, *oriT* RK2	Bartosik et al. (1997)
pABW3	Km^r, *ori* pMB1, *ori* pTAV202, mobilizable cloning vector, *lacZα*, *oriT* RK2	Bartosik et al. (2001)
pABW3-hipAB1	pABW3 derivative containing *hipAB* module of pKON1 (amplified by PCR with primers LHIPSBX and RHIPABS) cloned between XbaI and SacI sites of MCS	This study
pABW3-hipAB2	pABW3 derivative containing the *hipAB* module of *P. aminophilus* JCM 7686 chromosome (amplified by PCR with primers LHIPAPHS and RHIPAPHX) cloned between XbaI and SacI sites of MCS	This study
pABW3-hipAB3	pABW3 derivative containing the *hipAB* module of the plasmid pAMI8 of *P. aminophilus* JCM 7686 (amplified by PCR with primers LHIPAMI8 and RHIPAMI8) cloned between XbaI and EcoRI sites of MCS	This study
pJCB1	Km^r, *ori* pSC101, *ori* pKON1, mobilizable cloning shuttle vector, *lacZα*, *oriT* RK2	This study
pJCB1-hipAB1	Derivative of pJCB1 containing the *hipAB* module of pKON1 (amplified by PCR with primers LHIPSBX and RHIPABS) cloned between XbaI and SacI sites of MCS	This study
pJCB1-ΔhipA	pJCB1-hipAB1 with 893-bp deletion within *hipA* gene, generated by digestion with AgeI and BstBI	This study
pABW2-hipA	pABW2 derivative containing internal part of *hipA* gene of pKON1 (amplified by PCR with primers LHIPA and RHIPAHI) cloned between EcoRI and HindIII sites of MCS	This study
pRK2013	Km^r, helper plasmid carrying genes for conjugal transfer from RK2	Ditta et al. (1980)

2.2. DNA manipulations and PCR conditions

Plasmids of *Paracoccus* spp. were isolated according to the method of Birnboim and Doly (1979), and when required, the DNA was further purified by CsCl–ethidium bromide gradient centrifugation (Sambrook and Russell, 2001). Plasmid DNA was isolated from *E. coli* cells using a Plasmid Mini Kit (A&A Biotechnology). DNA manipulations were performed according to standard methods described by Sambrook and Russell (2001). Restriction endonucleases and T4 DNA ligase were used according to the supplier's instructions (Thermo Scientific). Amplification by PCR was performed in a Mastercycler (Eppendorf) using synthetic oligonucleotide primers (listed in Supplementary Table S1), Phusion polymerase (Thermo Scientific) and appropriate DNA templates. The PCR-amplified DNA fragments were analyzed by electrophoresis on 0.8% agarose gels and, where necessary, purified using a Gel Out Kit (A&A Biotechnology).

2.3. Introduction of DNA into bacterial cells

Chemical transformation of *E. coli* strains was performed according to the method of Kushner (1978). Triparental mating to introduce plasmids into *Paracoccus* spp. cells was performed as described previously (Bartosik et al., 2001). Briefly, for triparental mating, the donor strain *E. coli* TG1 carrying a mobilizable vector, *E. coli* DH5α carrying the helper plasmid pRK2013 (Ditta et al., 1980), and a suitable recipient strain (*P. kondratievae* NCIBM 131773R, *P. kondratievae* UW300, *P. aminophilus* JCM 7686R, *P. aminophilus* UW200, *P. versutus* UW225 or *P. pantotrophus* KL100) were mixed at a ratio of 1:1:2. One hundred microliters of this mixture was spread onto a plate containing solidified LB medium. After overnight incubation at 30 °C, the bacteria were washed off the plate, and suitable dilutions were plated on selective media containing rifampin (selectable markers of the recipient strains) and kanamycin to select for transconjugants. The plasmid content of transconjugants was verified by screening several colonies using a rapid alkaline extraction procedure and agarose gel electrophoresis. Spontaneous resistance of the recipient strains to the antibiotics used for selection was not detectable under these experimental conditions.

2.4. Plasmid stability assay

The stability of shuttle plasmids (pJCB1, pJCB1-hipAB1, pJCB1-ΔhipA, pABW3, pABW3-hipAB1, pABW3-hipAB2, pABW3-hipAB3) in *Paracoccus* spp. strains was tested during growth under nonselective conditions as described previously (Dziewit et al., 2007). Briefly, stationary-phase cultures were diluted in fresh medium without antibiotic selection and cultivated for approximately 10, 20, and 30 generations. Samples were diluted and plated onto solid medium in the absence of selective antibiotics. One hundred colonies were tested for the presence of the plasmid-encoded Km^r markers by replica plating. Plasmid retention was determined by using the percentage of kanamycin resistant colonies.

2.5. DNA sequencing and bioinformatic analysis

The complete nucleotide sequence of plasmid pKON1 was determined by the DNA Sequencing and Oligonucleotide Synthesis Laboratory (oligo.pl) at the Institute of Biochemistry and Biophysics, Polish Academy of Sciences, Warsaw. High-throughput sequencing of a MID-tagged shotgun plasmid-library was performed using a FLX Titanium Genome Sequencer (Roche/454 Life Sciences). Newbler de novo assembler software (Roche) was used for the sequence assembly. Final gap closure and sequence polishing were performed by capillary sequencing of PCR products using an ABI3730xl DNA Analyzer (Applied Biosystems). The plasmid nucleotide sequence was analyzed using Clone Manager (Sci-Ed8) and Artemis software (Carver et al., 2008). Similarity searches were performed using the BLAST programs (Altschul et al., 1997) provided by the National Center for Biotechnology Information (NCBI) (http://blast.ncbi.nlm.nih.gov/Blast.cgi) and the PRIAM tool (Claudel-Renard et al., 2003). Comparison searches of insertion sequences were performed with ISfinder (Siguier et al., 2006). Promoter prediction was performed using BPROM software (Solovyev and Salamov, 2011).

2.6. Nucleotide sequence accession number

The annotated nucleotide sequence of plasmid pKON1 has been deposited in the GenBank database with the accession number KP294352. The nucleotide sequences of insertion sequences identified in this study have been submitted to the ISfinder database.

3. Results and discussion

3.1. Identification of pKON1 as a putative essential replicon

Bacteria of the genus *Paracoccus* frequently carry large extrachromosomal replicons, some of which (chromids) contain genes that are essential for the viability of the host strain (Dziewit et al., 2014). The initial goal of this study was to identify chromids among large replicons (approx. 70–150 kb) of three strains of this genus: *P. kondratievae* NCIMB 13773^T, *P. homiensis* DSM 11862, and *P. aminovorans* JCM 7685. In a preliminary study, the extraction of plasmid DNA and DNA electrophoresis revealed that these strains contain four large replicons, designated (i) pKON1 (95 kb; strain NCIMB 13773^T), (ii and iii) pHOM1 and pHOM4 (100 kb and 140 kb, respectively; DSM 11862) and (iv) pAMV1 (185 kb; JCM 7685) (DSM 11862 and JCM 7685 contain also much smaller plasmids of the sizes 6 kb and 4.2 kb, respectively) (data not shown).

To test whether these replicons are essential for their hosts, the target-oriented replicon curing technique based on incompatibility was used, *i.e.* the inability of two plasmids carrying related replication systems to stably coexist in a bacterial cell (Dziewit and Bartosik, 2015). Using this strategy, we found that plasmids pHOM1, pHOM4 and pAMV1 were readily removed from the cells, while all attempts to remove pKON1 failed. Only a few *P. kondratievae* transconiugants were obtained, but none of them contained

autonomous forms of the introduced shuttle plasmid carrying replication system of pKON1.

These data suggested that pKON1 is an indispensable replicon or, alternatively, that it contains an efficient maintenance system, precluding removal of the parental replicon from the host cells. To investigate the nature of the observed phenomenon, the genomic structure of pKON1 was examined.

3.2. Structure and genetic load of pKON1

The complete nucleotide sequence of plasmid pKON1 was determined. Sequence analysis revealed that pKON1 is a circular molecule of 95,049 bp in size, whose structure is unique among plasmids sequenced to date (Fig. 1).

The average GC content of the pKON1 sequence is 60.2%, which is close to the value determined for the chromosomal DNA of the host strain *P. kondratievae* NCIMB 13773^T (62.5%; Doronina and Trotsenko, 2000). No "plasmid islets" with a significantly different GC content were detected in the pKON1 genome. Bioinformatic analysis revealed that pKON1 contains 97 predicted coding sequences (CDS), constituting 81.7% of the plasmid genome. Characteristics of these putative genes, including their position, transcriptional orientation, the size of the encoded proteins, and their closest known homologs are shown in Supplementary Table S2.

Comparative analysis of the amino acid (aa) sequences of the predicted proteins permitted the identification of the pKON1 maintenance region, composed of replication (REP) and partitioning (PAR) systems. This region is highly conserved in several large alphaproteobacterial replicons, including (i) the chromid pAMI6 of *P. aminophilus* JCM 7686 (Dziewit et al., 2014), (ii) pTAV3, a mega-sized replicon of *P. versutus* UW1 (Bartosik et al., 2002a) and (iii) chromosome 2 of *Rhodobacter sphaeroides* 2.4.1 (Kontur et al., 2012). Plasmid pKON1 also carries a pair of genes encoding a putative type II TA system of the *hipAB* family (pKON1_p62-pKON1_p63), which is located distantly from the REP and PAR modules in the pKON1 genome (Fig. 1A and B). An identical *locus*, encoding a predicted HipB antitoxin and HipA toxin,is present in an 8640-bp sequence contig of *Paracoccus* sp. TRP (accession no. NZ_AEPN01000073). Closely

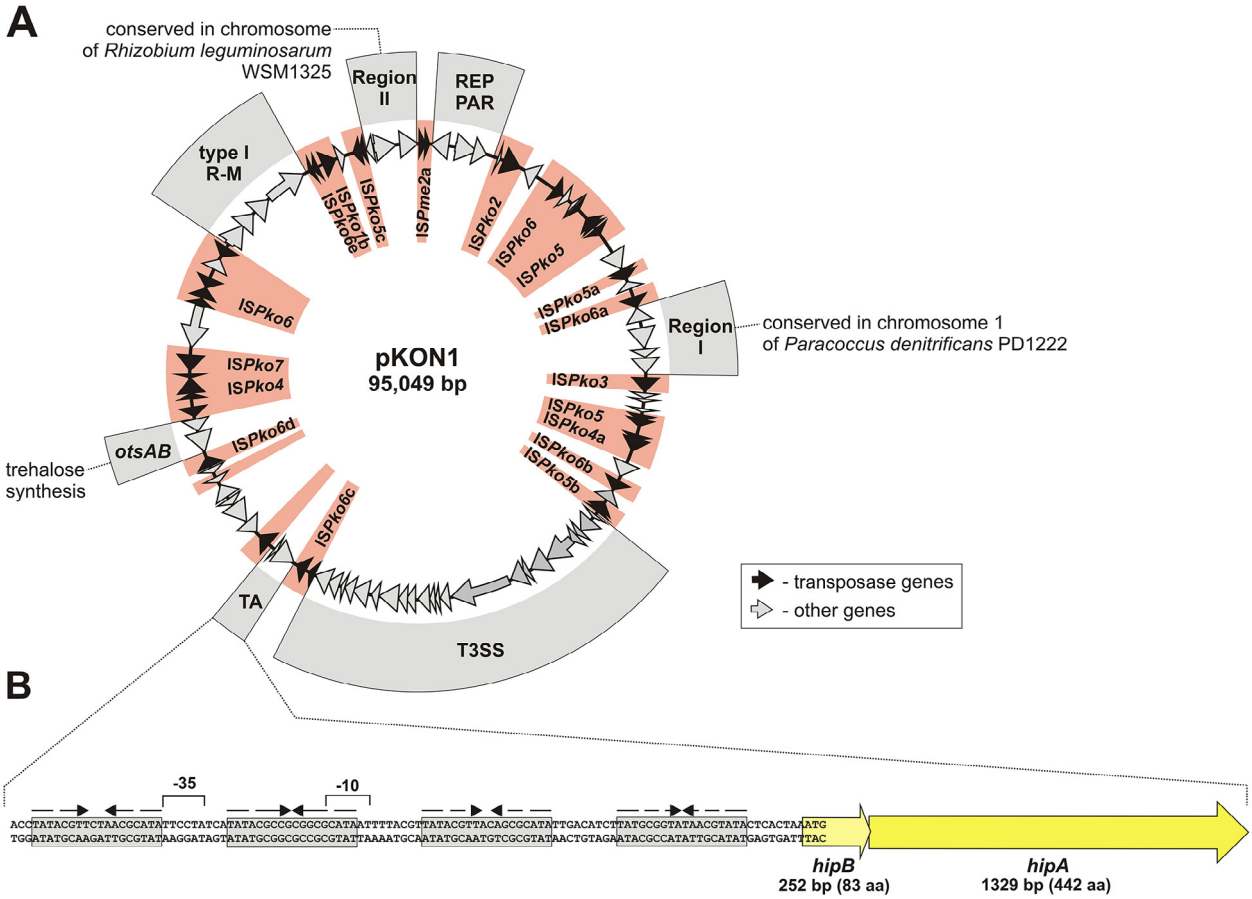


Fig. 1. Genetic structure of plasmid pKON1. (Panel A) Circular map of pKON1 comprising all predicted genes. The inner shaded blocks indicate DNA regions influenced by transposition. The location of complete ISs is marked. The outer part of the map shows conserved DNA regions and predicted genetic modules of pKON1. (Panel B) Genetic organization of the TA module of pKON1. The large shaded arrows represent the genes for the antidote (*hipB*) and the toxin (*hipA*), and indicate their transcriptional orientation (the gene length and estimated size of the predicted proteins are given below the arrows). The nucleotide sequence of the DNA region containing the *hipAB* promoter is shown and putative –35 and –10 motifs (BPROM software) are marked. Predicted sites of interaction with the HipB antitoxin are boxed and shaded, and imperfect palindromic sequences are denoted by dashed arrows.

related gene pairs are also found within sequence contigs of *Sulfitobacter* sp. H3 and *Oceanibulbus indolifex* HEL-45 (accession nos. JAMD01000016 and ABID01000075, respectively).

It was apparent that the structure of pKON1 has been significantly shaped by transposition, since almost 30% of the plasmid genome is composed of complete (20 elements) or incomplete (16) insertion sequences (ISs) (Fig. 1). Based on detailed comparative sequence analysis performed in the ISfinder database (Siguier et al., 2006), the majority of the predicted complete elements were classified into five IS families: (i) IS*3* (IS*150* group), (ii) IS*5* (IS*427* and IS*903* groups), (iii) IS*6*, (iv) IS*110* (IS*1111* group) and (v) IS*256* (Table 2). Most of the remnant ISs displayed highest similarities to the aforementioned intact elements. However, a few were similar to the 3′- or 5′-end parts of ISs classified to other recognized families/groups: (i) IS*3* (IS*3* and IS*51* groups), (ii) IS*66* and (iii) IS*701* (see Supplementary Table S3).

Six of the outlined ISs are novel elements, which have been given the following designation: (i) IS*Pko2* (IS*256* family), (ii and iii) IS*Pko3* and IS*Pko5* (IS*5* family), (iv) IS*Pko4* (IS*3* family, IS*150* group), (v) IS*Pko6* (IS*6* family) and (vi) IS*Pko7* (IS*110* family) (Table 2). The remaining ISs were isoforms of elements identified previously in *Paracoccus* spp. strains by their transposition into trap plasmids: (i) IS*Pme2* (IS*5* family, IS*427* group) (Bartosik et al., 2008), (ii) IS*Ppa9* (IS*6* family) and (iii) IS*Pko1* (IS*5* family, IS*903* group) (Dziewit et al., 2012). Three of the ISs were present in multiple copies in the pKON1 genome: IS*Pko6*, IS*Pko5* and IS*Pko4* – 7, 5 and 2 copies, respectively (Fig. 1).

In a previous study, we identified functional TEs of *P. kondratievae* NCIMB 13773R by the use of trap plasmids, enabling positive selection of transposition events (Dziewit et al., 2012). However, this analysis did not result in the capture of any of the ISs carried by pKON1, which indicates that these elements are not functional or they transpose with a relatively low frequency. Many but not all of the ISs in the pKON1 genome are bordered by direct repeats (DRs), resulting from the duplication of the target sequence during transposition. DRs were observed in the case of (i) the elements of the IS*427* group (IS*Pme2*a and isoforms of IS*Pko5*; duplicated TA dinucleotide), (ii) IS*Ppa4*a, whose transposition disrupted a transposase gene of an insertion sequence IS*Pko6* and generated an 8-bp-long repeated sequence 5′-TTTCCTGG-3′ and (iii) IS*Pko2*, flanked by the 8-bp sequence 5′-GCGCCTTT-3′.

Interestingly, elements IS*Pme2*a and IS*Pko5*a have been inserted within the stop codons of genes encoding a replication initiator protein RepB (pKON1_p1) and a conserved protein with sequence similarity to the 50S ribosomal protein L20 (pKON1_p16), respectively. In both cases, the disrupted codons were replaced by a TAG fusion stop codon, generated by target site duplication (TA) and IS insertion (G being the first nucleotide of the ISs). These unique insertion events allowed the expression of the authentic proteins.

Plasmid pKON1 also contains additional genetic load comprising numerous genes commonly found in many bacterial chromosomes, including a highly conserved type III secretion system (T3SS) and a type I restriction–modification system (R–M). The T3SS module constitutes approximately 21% of the pKON1 genome (nucleotide position 34,045–54,578) and is composed of 22 genes (pKON1_p38–pKON1_p59) placed in the same transcriptional orientation. These genes encode homologs of all the core proteins of the well-defined T3SS systems, including (i) ATPase SctN (pKON1_p53), (ii) a negative regulator SctL (pKON1_p52), (iii and iv) T3SS gate proteins SctU (pKON1_p59) and SctV (pKON1_p42), (v) SctJ, a protein forming the inner ring of the basal body (pKON1_p50), (vi) SctQ, a protein forming the cytoplasmic ring (pKON1_p52) and (vii–ix) three core membrane proteins SctR (pKON1_p56), SctS (pKON1_p57) and SctT (pKON1_p58). The most closely related gene cluster was identified in the chromosome of *Paracoccus aminophilus* JCM 7686 (accession no. CP006650; nucleotide position 1,791,291–1,807,830). Based on the T3SS module structure and sequence similarities, we classified the predicted secretion system of pKON1 as a member of subgroup II of the *Rhizobiales*-T3SS family (Rhc-T3SS) (Gazi et al., 2012).

Another conserved gene cluster of pKON1 consists of four genes (nucleotide position 80,262–87,304) encoding proteins with significant sequence similarities to components of a type IC R–M system (Murray, 2000): HsdM (pKON1_p85), HsdS (pKON1_p86), PrrC (pKON1_p87) and HsdR (pKON1_p88). The most closely related proteins are encoded by the chromosome of *Bordetella avium* 197N (EcoR124II; accession no. AM167904; nucleotide position 564,642–571,720). The HsdR, HsdM and PrrC proteins of both systems show aa sequence identity in the range 89–92%, whereas the HsdS subunits (responsible for the specificity of sequence recognition) are much less well conserved (40% identity).

It has been demonstrated that the HsdM, HsdS and HsdR subunits form multifunctional protein complexes with methyltransferase or restriction activity (Murray, 2000), while the PrrC protein is a phage-excluding, anti-codon nuclease (ACNase), able to cleave phage tRNALys (Blanga-Kanfi et al., 2006). Co-localization of *prrC* with the *hsdMSR* gene cluster, first observed in the EcoprrI R–M system (Tyndall et al., 1994), is conserved in diverse bacteria, suggesting some functional coupling.

Plasmid pKON1 also contains two putative genes (*ostA* and *ostB*), encoding predicted trehalose-6-phosphate synthase (OstA; pKON1_p72) and trehalose-6-phosphate phosphatase (OstB; pKON1_p73). These enzymes (conserved

Table 2
Transposable elements identified in the pKON1 genome.

IS (and its isoforms)	IS family/group	Length (bp)	IR (bp)[a]	DR (bp)
IS*Pko6* (IS*Pko6*a–e)	IS*6*	817	17/13	–
IS*Ppa9*a	IS*6*	812	15/15	–
IS*Pko5* (IS*Pko5*a–c)	IS*5*/ IS*427*	851	15/12	2
IS*Pme2*a	IS*5*/ IS*427*	851	15/13	2
IS*Pko3*	IS*5*/IS*903*	1050	17/16	–
IS*Pko1*b	IS*5*/IS*903*	1050	18/18	–
IS*Pko4* (IS*Pko4*a)	IS*3*/IS*150*	1524	30/21	8
IS*Pko2*	IS*256*	1396	37/29	8
IS*Pko7*	IS*110*/IS*1111*	1140	12/11	–

[a] The length of the IRs/the number of conserved residues.

among bacteria, fungi, plants and insects) are responsible for the synthesis of trehalose from UDP-glucose and D-glucose-6-phosphate (Avonce et al., 2006). The putative OstA and OstB enzymes encoded by pKON1 are most closely related to proteins encoded within genomes of the *Rhizobiales*, *e.g. Sinorhizobium fredii* HH103 (accession nos. YP_005189888 and YP_005189889), displaying 70% and 49% aa sequence identity, respectively.

Two other DNA regions carrying genes that are highly conserved (in sequence and synteny) in the chromosomes of other bacteria were also found within the pKON1 genome. The first gene cluster, localized between IS*Pko6*a and IS*Pko3*, is conserved in chromosome 1 of *Paracoccus denitrificans* PD1222 (accession no. NC_008686; nucleotide position 1,077,056–1,081,691; 99% nucleotide sequence identity). This region contains four genes encoding putative proteins: (i) a transcription regulator of the LysR family (pKON1_p21), (ii) a major facilitator superfamily transporter (pKON1_p22), (iii) 4-vinyl reductase (pKON1_p23) and (iv) peptidase M19 (pKON1_p24) (Fig. 1). The second gene cluster of pKON1 consists of three putative genes (pKON1_p71-pKON1_p73) (Fig. 1), whose homologs are localized in synteny in the chromosome of *Rhizobium leguminosarum* bv. trifolii WSM1325 (accession no. CP001622; nucleotide position 1,843,584–1,847,080). These genes encode the following putative proteins: (i) MarR family transcriptional regulator (pKON1_p95), (ii) FAD-binding monooxygenase (pKON1_p96) and (iii) a multidrug efflux protein (pKON1_p97). Interestingly, the first two genes of this cluster are also conserved in chromosome 2 of *Paracoccus denitrificans* PD1222 (accession no. NC_008687; nucleotide position 225,122–227,174). The specific role of the proteins encoded by the genes occurring within both DNA regions remains unknown.

In summary, our bioinformatic analysis revealed that pKON1 has a patchwork structure and contains numerous genes found mostly in the chromosomes of bacteria of the order *Rhizobiales* (Table S2). This finding is not surprising since the host strain of pKON1, *P. kondratievae* NCIMB 13773^T, was isolated from the maize rhizosphere (Doronina and Trotsenko, 2000), which is the natural habitat of rhizobia.

The presence of numerous IS elements, both complete and remnant, strongly suggests that transposition was the main force responsible for the observed gene shuffling in pKON1. Moreover, the identification of partial genes adjacent to several ISs and the failure to observe any continuation of these genes on the other side of the elements suggest that homologous recombination between copies of the ISs located in different replicons might also have occurred.

The presence of a rhizobial type T3SS within pKON1 is particularly noteworthy. These secretion systems are commonly encoded by both plant pathogenic and symbiotic bacteria, which points to their crucial role in mediating diverse plant–bacteria interactions (Tampakaki, 2014).

3.3. Identification of the pKON1 stability determinant

To identify genes responsible for the high stability (or indispensability) of pKON1 in *P. kondratievae* cells, a novel mobilizable shuttle cloning vector pJCB1 was constructed (Table 1). pJCB1 contains two replication systems – one originating from pKON1 and the second from low copy number plasmid pSC1 (specific for *E. coli*, non-functional in *Paracoccus* spp.) (Table 1).

This vector was used to clone NheI restriction fragments of pKON1 in *E. coli* TG1. A pool of the obtained TG1 transformants containing pJCB1 recombinant derivatives served as the donor in a triparental mating with *P. kondratievae* NCIMB 13773R recipient cells. As a result, numerous transconiugants were obtained – all lacking pKON1. In each case, the loss of the plasmid was associated with the presence of the pJCB1 derivative containing a 6355-bp-long NheI restriction fragment of pKON1. This observation strongly suggested that the cloned DNA fragment, containing the *hipAB* family TA system, could compensate for the loss of pKON1.

To investigate the role of the TA system in the maintenance of pKON1, two additional pJCB1 derivatives were constructed: (i) pJCB1-hipAB1, carrying a DNA fragment (1872 bp) containing the *hipAB* system and (ii) pJCB1-ΔhipA, a derivative carrying the intact *hipB* gene and mutated *hipA* (deletion of a 886-bp-long internal fragment of *hipA*; the entire gene has 1320 bp) (Table 1). These plasmids were efficiently introduced via conjugation into *P. kondratievae* cells and in both cases all tested transconiugants were deprived of pKON1. This showed that the *hipB* gene alone (the source of antitoxin that counteracts the toxin) could compensate for the loss of pKON1 from bacterial cells. The role and activity of the *hipAB* system were additionally verified by inactivation of the *hipA* toxin gene within the pKON1 genome (insertional inactivation by integration of pABW2-hipA into pKON1; Table 1). In contrast to parental pKON1, this mutated plasmid could be readily removed from NCIMB 13773R cells by the incompatible "empty" vector pJCB1 (data not shown). These findings confirmed that the high stability of pKON1 is dependent on the presence of the functional *hipAB* system. The predicted structural components of this system are shown in Fig. 1B.

3.4. Distribution of hipAB systems in bacterial genomes

NCBI database searches revealed that TA systems of the *hipAB* family are conserved and widely distributed in bacterial genomes. Homologous gene pairs were identified in both Gram-negative and Gram-positive bacteria (Leplae et al., 2011; data not shown). Related *loci* were also found in *Paracoccus* spp. genomes, *e.g. P. aminophilus* JCM 7686 carries two *hipAB loci*, in both the chromosome and the large plasmid pAMI8 (202 kb) (Dziewit et al., 2014), while *P. denitrificans* PD1222 (type strain of the genus *Paracoccus*) contains only one chromosomally-located *hipAB* locus (accession no. NC_008686; Pden_1548, Pden_1547).

The vast majority of the identified TA modules are present in bacterial chromosomes and they occur less frequently in plasmids and chromids. A few of the extrachromosomal replicons contain more than one copy of related *hipAB loci*, *e.g.* chromosome 2 (chromid) of *Variovorax paradoxus* S110 carries 3 systems of this type (accession no. NC_012792; Vapar_5917 and Vapar_5916, Vapar_6086 and Vapar_6085, Vapar_5936 and Vapar_5937) and pNGR234a of *Sinorhizobium fredii* NGR234 has 2 systems (accession no. NC_000914; NGR_a02540 and NGR_a02530, 2NGR_a04040 and NGR_a04050) (Dombrecht et al., 2001).

The archetype module of the *hipAB* family was identified in the chromosome of *E. coli* (Black et al., 1991). It was shown that the activity of this module leads to inhibition of cell growth and to the development of a persister cell subpopulation (Correia et al., 2006; Germain et al., 2013). Far less is known about plasmid-encoded *hipAB* systems. To our knowledge, only two such *loci* have been analyzed to date: *spsAB* and *spsCD*, both from the aforementioned plasmid pNGR234a (HipA of pKON1 shows only 28% and 35% amino acid sequence identity, respectively, with SpsB and SpsD HipA-like proteins). These TA systems were shown to significantly improve stability of a tested plasmid, although they did not completely prevent the appearance of plasmid-less cells in the bacterial population (Dombrecht et al., 2001).

3.5. *Stabilization properties of hipAB loci of* Paracoccus *spp.*

Preliminary stabilization assays with pJCB1-hipAB1 (contains REP, PAR and TA modules of pKON1) revealed that this plasmid could be very stably maintained in *P. kondratievae* UW300 (pKON1-less strain), since no plasmid-less cells were detected after approximately 30 generations of growth under nonselective conditions. In contrast, the plasmid pJCB1-ΔhipA (with a mutated *hipA* gene) and parent vector pJCB1 (containing only the REP and PAR regions of pKON1) remained in approximately 65% of bacterial cells tested under analogous conditions (see Supplementary Fig. S1).

The ability of the *hipAB* system to stabilize a heterologous replicon was then tested. This analysis was extended to include two related TA systems originating from plasmid pAMI8 and the chromosome of *P. aminophilus* JCM 7686 [accession nos. NC_022050 (JCM7686_pAMI8p105, JCM7686_pAMI8p106) and NC_022041 (JCM7686_1676, hipBJCM7686_1677), respectively)]. PCR-amplified DNA fragments containing these TA systems were cloned into the low-copy-number shuttle vector pABW3, which is unstable in paracocci (Km^r; carries the heterologous replication system of plasmid pTAV1 of *P. versutus* UW1). The three resulting plasmids, pABW3-hipAB1 (contains the TA module of pKON1), pABW3-hipAB2 (TA of JCM 7686 chromosome) and pABW3-hipAB3 (TA of pAMI8) (Table 1) were introduced separately into strains representing four *Paracoccus* species: (i) *P. kondratievae* NCIBM 131773R (in the case of pABW3-hipAB1, pKON1-less strain UW300 was used), (ii) *P. aminophilus* JCM 7686R (in the case of pABW3-hipAB3, pAMI8-less strain UW200 was used), (iii) *P. versutus* UW225 (pTAV1-less) and (iv) *P. pantotrophus* KL100. The stability of the introduced plasmids was then determined.

The *hipAB* modules originating from pAMI8 and pKON1 were found to promote stability of pABW3-derivatives under non-selective conditions in all tested hosts, although they differed in their stabilization properties (Fig. 2). The stabilization effect conferred by the former module was host dependent, with significantly lower plasmid stability observed in *P. pantotrophus* KL100 (Fig. 2C). The module of pKON1 was highly efficient and its activity precluded the detection of any plasmid-less cells (Fig. 2), even after approximately 100 generations of growth in continuous culture (data not shown). In contrast, the module of chromosomal origin was unable to stabilize the test plasmid in any of the hosts.

4. Conclusions

The analysis performed in this study revealed that pKON1 is not an indispensable replicon, therefore it should be considered as a plasmid rather than a chromid. Since the essential nature of the predicted chromids is usually verified by the application of a target-oriented replicon-curing technique, it is very important to identify the genes underlying the apparent "essentiality" (*e.g.* efficient TA or R–M systems) before labeling a replicon as a chromid or plasmid.

We showed that the high stability of plasmid pKON1, demonstrated by the failure of attempts to remove this plasmid from *P. kondratievae* cells, is determined by the presence and activity of a highly efficient *hipAB* family TA system. Such a high level of stabilization has not been previously observed in large plasmids of *Paracoccus* spp. For example, plasmid pAMI8 of *P. aminophilus* JCM 7686 can be readily removed from JCM 7686 cells by incompatibility, despite the presence of four TA modules representing the *relBE*/*parDE* (2 systems) and *hipAB* families (the latter TA module was tested in this study), as well as a hybrid module *phd/vapC*. Similar behavior was observed in the case of two other replicons of the JCM 7686 strain, pAMI4 (438 kb) and pAMI6 (206 kb), containing *phd/doc* and *ccdAB* family systems, respectively, as well as nine other large plasmids of *P. pantotrophus*, *P. versutus*, *P. aminovorans*, *P. homiensis* and *P. solventivorans*, whose complete nucleotide sequences have yet to be determined (Bartosik et al., 1998, 2002a, 2002b; this study and unpublished results).

We also showed that the high level of stability provided by the *hipAB* module of pKON1 is not replicon or host dependent. Therefore this TA system is an excellent candidate for the construction of a stabilization cassette, useful in the creation of shuttle vectors specific for *Paracoccus* spp. and possibly the entire *Alphaproteobacteria*.

The *hipAB* family systems are mainly found in bacterial chromosomes. There has been considerable interest in these *loci* since they may be involved in the production of persister cells within bacterial populations. The *E. coli* HipA toxin is a kinase that inhibits glutamyl-tRNA synthetase (GltX) by phosphorylation. Such inhibition results in the accumulation of uncharged tRNAGlu and in increased synthesis of the alarmone ppGpp, which significantly increases the level of persistence (Germain et al., 2013).

It will be of great interest to identify the cellular targets of toxins encoded by the *hipAB* module of pKON1 and the other TA systems examined in this study. Differences in the stabilization properties of these systems suggest that their toxins may have distinct cellular targets. Such analysis will demonstrate whether plasmid-encoded *hipAB* modules may serve as portable systems enabling the formation of dormant persister cells that are highly resistant to environmental stress conditions.

Acknowledgments

This work was supported by the National Science Centre in Poland (grant no 2013/09/B/NZ1/00133).

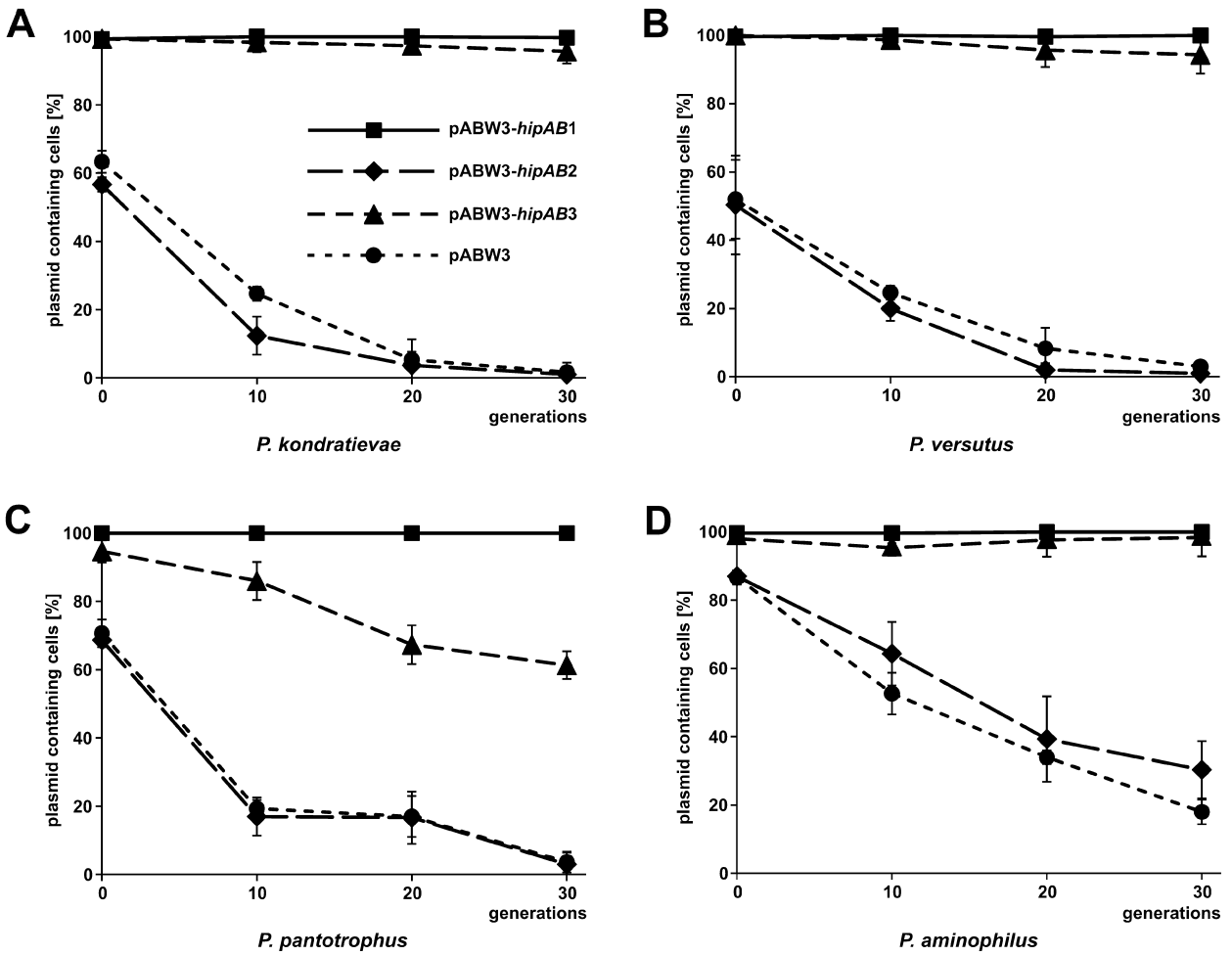


Fig. 2. Stability of stability probe vector pABW3 and derivatives containing TA modules of pKON1 of *P. kondratievae* NCIMB 13773^T (pABW3-hipAB1), and the chromosome (pABW3-hipAB2) and plasmid pAMI8 (pABW3-hipAB3) of *P. aminophilus* JCM 7686, in different strains of *Paracoccus* spp. Plasmid stability was tested in *P. kondratievae* NCIMB 13773R (or pKON1-less strain UW300 in the case of pABW3-hipAB1) (panel A), *P. versutus* UW225 (panel B), *P. pantotrophus* KL100 (panel C) and *P. aminophilus* JCM 7686R (or pAMI8-less derivative strain UW200 in the case of pABW3-hipAB3) (panel D). The cultures were grown under non-selective conditions for approximately 10, 20 and 30 generations. Error bars denote standard deviations.

Appendix: Supplementary material

Supplementary data to this article can be found online at doi:10.1016/j.plasmid.2015.02.003.

References

Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W., et al., 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25, 3389–3402.

Avonce, N., Mendoza-Vargas, A., Morett, E., Iturriaga, G., 2006. Insights on the evolution of trehalose biosynthesis. BMC Evol. Biol. 6, 109.

Bartosik, D., Baj, J., Plasota, M., Piechucka, E., Wlodarczyk, M., 1993. Analysis of *Thiobacillus versutus* pTAV1 plasmid functions. Acta Microbiol. Pol. 39, 5–11.

Bartosik, D., Białkowska, A., Baj, J., Włodarczyk, M., 1997. Construction of mobilizable cloning vectors derived from pBGS18 and their application for analysis of replicator region of a pTAV202 mini-derivative of *Paracoccus versutus* pTAV1 plasmid. Acta Microbiol. Pol. 46, 387–392.

Bartosik, D., Baj, J., Wlodarczyk, M., 1998. Molecular and functional analysis of pTAV320, a *repABC*-type replicon of the *Paracoccus versutus* composite plasmid pTAV1. Microbiology 144, 3149–3157.

Bartosik, D., Szymanik, M., Wysocka, E., 2001. Identification of the partitioning site within the repABC-type replicon of the composite *Paracoccus versutus* plasmid pTAV1. J. Bacteriol. 183, 6234–6243.

Bartosik, D., Baj, J., Bartosik, A.A., Wlodarczyk, M., 2002a. Characterization of the replicator region of megaplasmid pTAV3 of *Paracoccus versutus* and search for plasmid-encoded traits. Microbiology 148, 871–881.

Bartosik, D., Baj, J., Piechucka, E., Waker, E., Wlodarczyk, M., 2002b. Comparative characterization of *repABC*-type replicons of *Paracoccus pantotrophus* composite plasmids. Plasmid 48, 130–141.

Bartosik, D., Putyrski, M., Dziewit, L., Malewska, E., Szymanik, M., Jagiello, E., et al., 2008. Transposable modules generated by a single copy of insertion sequence IS*Pme1* and their influence on structure and evolution of natural plasmids of *Paracoccus methylutens* DM12. J. Bacteriol. 90, 3306–3313.

Birnboim, H.C., Doly, J., 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res. 7, 1513–1523.

Black, D.S., Kelly, A.J., Mardis, M.J., Moyed, H.S., 1991. Structure and organization of hip, an operon that affects lethality due to inhibition of peptidoglycan or DNA synthesis. J. Bacteriol. 173, 5732–5739.

Blanga-Kanfi, S., Amitsur, M., Azem, A., Kaufmann, G., 2006. PrrC-anticodon nuclease: functional organization of a prototypical bacterial restriction RNase. Nucleic Acids Res. 34, 3209–3219.

Carver, T., Berriman, M., Tivey, A., Patel, C., Böhme, U., Barrell, B.G., et al., 2008. Artemis and ACT: viewing, annotating and comparing sequences stored in a relational database. Bioinformatics 24, 2672–2676.
Claudel-Renard, C., Chevalet, C., Faraut, T., Kahn, D., 2003. Enzyme-specific profiles for genome annotation: PRIAM. Nucleic Acids Res. 31, 6633–6639.
Correia, F.F., D'Onofrio, A., Rejtar, T., Li, L., Karger, B.L., Makarova, K., et al., 2006. Kinase activity of overexpressed HipA is required for growth arrest and multidrug tolerance in *Escherichia coli*. J. Bacteriol. 188, 8360–8367.
Ditta, G., Stanfield, S., Corbin, D., Helinski, D.R., 1980. Broad host-range DNA cloning system for Gram-negative bacteria: construction of a bank of *Rhizobium meliloti*. Proc. Natl. Acad. Sci. U.S.A. 77, 7347–7351.
Dombrecht, B., Vanderleyden, J., Michiels, J., 2001. Stable RK2-derived cloning vectors for the analysis of gene expression and gene function in gram-negative bacteria. Mol. Plant Microbe Interact. 14, 426–430.
Doronina, N.V., Trotsenko, I.A., 2000. A novel plant-associated thermotolerant alkalophilic methylotroph of the genus *Paracoccus*. Microbiology 69, 593–598.
Dziewit, L., Bartosik, D., 2015. Comparative analyses of extrachromosomal bacterial replicons, identification of chromids, and experimental evaluation of their indispensability. Methods Mol. Biol. 1231, 15–29.
Dziewit, L., Jazurek, M., Drewniak, L., Baj, J., Bartosik, D., 2007. The SXT conjugative element and linear prophage N15 encode toxin-antitoxin-stabilizing systems homologous to the *tad-ata* module of the *Paracoccus aminophilus* plasmid pAMI2. J. Bacteriol. 189, 1983–1997.
Dziewit, L., Baj, J., Szuplewska, M., Maj, A., Tabin, M., Czyzkowska, A., et al., 2012. Insights into the transposable mobilome of *Paracoccus* spp. (*Alphaproteobacteria*). PLoS ONE 7, e32277.
Dziewit, L., Czarnecki, J., Wibberg, D., Radlinska, M., Mrozek, P., Szymczak, M., et al., 2014. Architecture and functions of a multipartite genome of the methylotrophic bacterium *Paracoccus aminophilus* JCM 7686, containing primary and secondary chromids. BMC Genomics 15, 124.
Gazi, A.D., Sarris, P.F., Fadouloglou, V.E., Charova, S.N., Mathioudakis, N., Panopoulos, N.J., et al., 2012. Phylogenetic analysis of a gene cluster encoding an additional, rhizobial-like type III secretion system that is narrowly distributed among *Pseudomonas syringae* strains. BMC Microbiol. 12, 188.
Germain, E., Castro-Roa, D., Zenkin, N., Gerdes, K., 2013. Molecular mechanism of bacterial persistence by HipA. Mol. Cell 52, 248–254.
Gibson, T.J., 1984. Studies on Epstein-Barr genome (Ph.D. thesis). University of Cambridge.
Goeders, N., Van Melderen, L., 2014. Toxin-antitoxin systems as multilevel interaction systems. Toxins 6, 304–324.
Hanahan, D., 1983. Studies on transformation of *Escherichia coli* with plasmids. J. Mol. Biol. 166, 557–580.
Harrison, P.W., Lower, R.P., Kim, N.K., Young, J.P., 2010. Introducing the bacterial 'chromid': not a chromosome, not a plasmid. Trends Microbiol. 18, 141–148.
Hayes, F., 2003. Toxins-antitoxins: plasmid maintenance, programmed cell death, and cell cycle arrest. Science 301, 1496–1499.
Kontur, W.S., Schackwitz, W.S., Ivanova, N., Martin, J., Labutti, K., Deshpande, S., et al., 2012. Revised sequence and annotation of the *Rhodobacter sphaeroides* 2.4.1 genome. J. Bacteriol. 194, 7016–7017.
Kushner, S.R., 1978. An improved method for transformation of *E. coli* with ColE1 derived plasmids. In: Boyer, H.B., Nicosia, S. (Eds.), Genetic Engineering. Elsevier/North Holland, Amsterdam, pp. 17–23.
Leplae, R., Geeraerts, D., Hallez, R., Guglielmini, J., Drèze, P., Van Melderen, L., 2011. Diversity of bacterial type II toxin–antitoxin systems: a comprehensive search and functional analysis of novel families. Nucleic Acids Res. 39, 5513–5525.
Maj, A., Dziewit, L., Czarnecki, J., Wlodarczyk, M., Baj, J., Skrzypczyk, G., et al., 2013. Plasmids of carotenoid-producing *Paracoccus* spp. (*Alphaproteobacteria*) – structure, diversity and evolution. PLoS ONE 8, e80258.
Murray, N.E., 2000. Type I restriction systems: sophisticated molecular machines (a legacy of Bertani and Weigle). Microbiol. Mol. Biol. Rev. 64, 412–434.
Pecota, D.C., Kim, C.S., Wu, K., Gerdes, K., Wood, T.K., 1997. Combining the *hok/sok*, *parDE*, and pnd postsegregational killer *loci* to enhance plasmid stability. Appl. Environ. Microbiol. 63, 1917–1924.
Sambrook, J., Russell, D.W., 2001. Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press, New York.
Siguier, P., Perochon, J., Lestrade, L., Mahillon, J., Chandler, M., 2006. ISfinder: the reference centre for bacterial insertion sequences. Nucleic Acids Res. 34, D32–D36.
Solovyev, V., Salamov, A., 2011. Automatic annotation of microbial genomes and metagenomic sequences. biomedicine and environmental studies. In: Li, R.W. (Ed.), Metagenomics and its Applications in Agriculture. Nova Science Publishers, pp. 61–78.
Stieber, D., Gabant, P., Szpirer, C., 2008. The art of selective killing: plasmid toxin/antitoxin systems and their technological applications. Biotechniques 45, 344–346.
Tampakaki, A.P., 2014. Commonalities and differences of T3SSs in rhizobia and plant pathogenic bacteria. Front. Plant Sci. 5, 114.
Tyndall, C., Meister, J., Bickle, T.A., 1994. The *Escherichia coli prr* region encodes a functional type IC DNA restriction system closely integrated with an anticodon nuclease gene. J. Mol. Biol. 237, 266–274.
Unterholzner, S.J., Poppenberger, B., Rozhon, W., 2013. Toxin-antitoxin systems: biology, identification, and application. Mob. Genet. Elements 3, e26219.

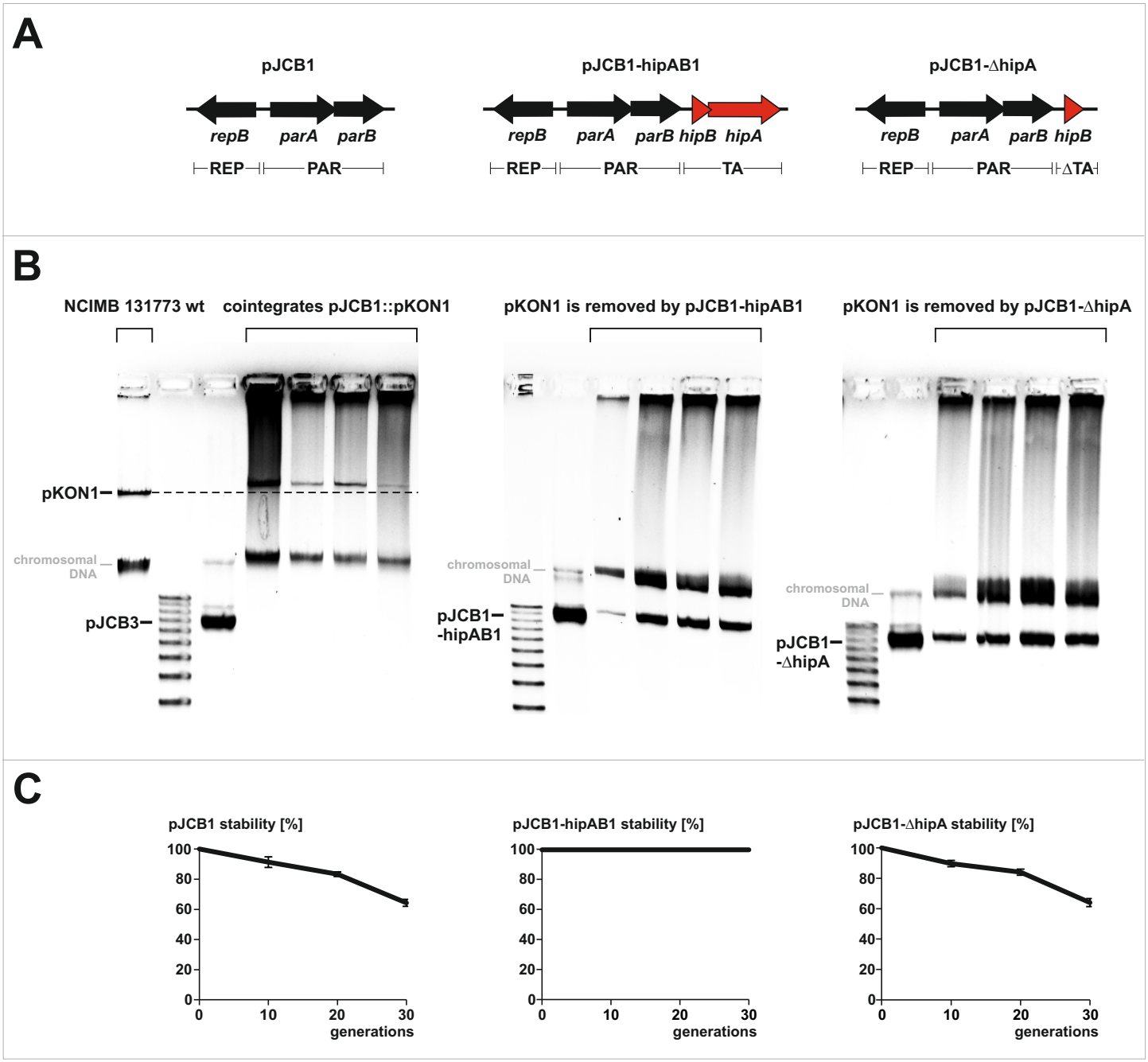


Fig. S1. Elimination of plasmid pKON1 from *P. kondratievae* NCIMB 13773R cells. (Panel A) REP, PAR and TA modules of pKON1 cloned in the shuttle plasmids pJCB1, pJCB-hipAB1 and pJCB1-ΔhipA. (Panel B) Plasmid profile of *P. kondratievae* transcojugants obtained after the introduction of the shuttle plasmids. (Panel C) Segregational stability of the shuttle plasmids in *P. kondratievae* UW300 (pKON1-less strain), tested under non-selective conditions.

Table S1.
Oligonucleotides used in this study.

No	Oligonucleotide	Sequence	Amplified region
1	LKON1EV	ATGATATCACATCACGTCGATCCGCCTC	replication and partition systems of pKON1 (used for pJCB1 construction)
2	RKON1EV	TAGATATCGATCTGCGCCTGTTCCTTGC	
3	LHIPSBX	CGTCTAGAATGGCCGAGACGACTGATGG	*hipAB* module of pKON1 (used for pABW3-hipA1 and pJCB1-hipAB1 construction)
4	RHIPABS	TAGAGCTCCATGGTGCCGGAAGAGTTCG	
5	LHIPA	ATCTGAACAACCGCCGCGTC	internal part of the *hipA* gene of pKON1 (used for pABW2-hipA construction)
6	RHIPAHI	CTAAGCTTACTCATTCTCAACGCTGTCC	
7	LHIPAPHS	CAGAGCTCGCCCTTTCTCAAGGTCTGTC	*hipAB* family system of *P. aminophilus* JCM 7686 *chromosome* (used for pABW3-hipA2 construction)
8	RHIPAPHX	CGTCTAGAGCAGAACCACCCTTTGTAAC	
9	LHIPAMI8	CGATGCCGCAATTGGATAAG	*hipAB* family system of pAMI8 (used for pABW3-hipA3 construction)
10	RHIPAMI8	TATCTAGACTTCACGCATCCTCTCAGAC	

Table S2.
Predicted genes of plasmid pKON1 (*Paracoccus kondratievae* NCIMB 13773^T).

Gene no.	Coding region (bp)	Strand	Protein size (aa)	Possible function	Best BLAST hits		
					% identity (aa)	Organism	GenBank accession no.
1	75-506	→	143	transposase OrfA (IS*Pme2*a)	99% (142/143)	*Paracoccus* sp. N5	WP_017998117
2	503-862	→	119	transposase OrfB (IS*Pme2*a)	100% (119/119)	*Paracoccus* sp. N5	WP_017998118
3	864-2042	←	392	replication protein	90% (355/394)	*Paracoccus pantotrophus*	WP_024845810
4	2499-3965	→	488	plasmid partitioning ATPase ParA	93% (455/488)	*Paracoccus versutus*	WP_032491260
5	3958-4926	→	322	plasmid partitioning protein ParB	83% (266/319)	*Paracoccus pantotrophus*	WP_024845808
6	5425-5610	→	61	hypothetical protein	58% (19/33)	*Paracoccus yeei*	WP_028718820
7	5841-7070	→	409	transposase (IS*Pko2*)	88% (361/409)	*Paracoccus aestuarii*	ADF28544
8	7349-8233	←	294	TonB-dependent receptor protein, Y4CG-like putative resolvase	95% (278/293)	*Rhizobium* sp. NT-26	CCF22029
9	9712-10395	→	227	transposase (IS*Pko6*)	99% (225/227)	*Paracoccus* sp. TRP	WP_010398086
10	10496-10906	→	136	hypothetical protein	98% (133/136)	*Rhodobacteraceae* bacterium PD-2	WP_023849205
11	11193-11441	←	82	truncated single-stranded DNA-binding protein	100% (82/82)	*Rhodobacteraceae*	WP_008335991
12	11546-11869	→	107	transposase OrfA (truncated copy of an IS)	99% (106/107)	*Paracoccus* sp. 39524	KGJ18543
13	12597-13145	←	182	transposase OrfB (IS*Pko5*)	87% (146/172)	*Paracoccus pantotrophus*	ACG76127
14	12947-13372	←	141	transposase OrfA (IS*Pko5*)	94% (132/141)	*Paracoccus pantotrophus*	WP_024844571
15	13539-14000	→	153	truncated transposase	100% (153/153)	*Paracoccus* sp. TRP	WP_010398086
16	15473-16690	→	405	50S ribosomal protein L20	55% (222/402)	*Leisingera methylohalidivorans*	WP_024091780
17	16757-17182	→	141	transposase OrfA (IS*Pko5*a)	95% (134/141)	*Paracoccus pantotrophus*	WP_024844571
18	16984-17532	→	182	transposase OrfB (IS*Pko5*a)	85% (146/171)	*Paracoccus pantotrophus*	ACG76127
19	17638-18141	←	167	hypothetical protein	46% (70/151)	*Bacteroides coprocola*	WP_007565586
20	18749-19432	→	227	transposase (IS*Pko6*a)	99% (226/227)	*Paracoccus* sp. TRP	WP_010398086
21	19514-20416	←	300	transcriptional regulator LysR family	99% (296/300)	*Paracoccus denitrificans*	WP_011747430

Gene no.	Coding region (bp)	Strand	Protein size (aa)	Possible function	Best BLAST hits		
					% identity (aa)	Organism	GenBank accession no.
22	20850-22208	→	452	major facilitator transporter	99% (450/452)	*Paracoccus denitrificans*	WP_011747431
23	22288-22821	→	177	4-vinyl reductase	98% (173/177)	*Paracoccus denitrificans*	WP_011747432
24	22879-23877	→	332	peptidase M19	99% (328/332)	*Paracoccus denitrificans*	WP_011747433
25	24185-25117	→	310	transposase (IS*Pko3*)	85% (262/309)	*Rubellimicrobium thermophilum*	WP_021099110
26	25194-25541	→	115	truncated chromosome partitioning protein ParB	99% (114/115)	*Paracoccus* sp. TRP	WP_010400476
27	25601-25834	→	77	hypothetical protein	92% (71/77)	*Paracoccus* sp. TRP	WP_010400474
28	26247-26483	→	78	hypothetical protein	100% (78/78)	*Paracoccus* sp. TRP	WP_010400470
29	26669-27094	→	141	transposase OrfA (IS*Pko5*)	94% (132/141)	*Paracoccus pantotrophus*	WP_024844571
30	26896-27444	→	182	transposase OrfB (IS*Pko5*)	87% (146/172)	*Paracoccus pantotrophus*	ACG76127
31	28027-28386	→	119	transposase OrfA (IS*Pko4*a)	100% (119/119)	*Paracoccus* sp. N5	WP_017998939
32	28383-29363	→	326	transposase OrfB (IS*Pko4*a)	94% (306/326)	*Sphingobium japonicum*	WP_006948954
33	30020-30997	→	325	transcriptional regulator AraC family	70% (225/322)	*Erwinia typographi*	KGT86037
34	31315-31998	→	227	transposase (IS*Pko6*b)	99% (224/227)	*Paracoccus* sp. TRP	WP_010398086
35	32090-32815	←	241	hypothetical protein	56% (124/223)	*Paracoccus aminophilus*	WP_020950502
36	33009-33611	←	200	transposase OrfB (IS*Pko5*b)	85% (146/172)	*Paracoccus pantotrophus*	WP_024844571
37	33532-33783	←	83	transposase OrfA (IS*Pko5*b)	95% (73/77)	*Paracoccus pantotrophus*	WP_024844571
38	34045-34284	→	79	type III secretion system component SctE	61% (46/76)	*Paracoccus* sp. 39524	KGJ11120
39	34508-35467	→	319	type III secretion system component SctD	39% (126/321)	*Paracoccus aminophilus*	WP_020950508
40	35479-35838	→	119	hypothetical protein	55% (49/89)	*Paracoccus aminophilus*	WP_020950507
41	35835-36362	→	175	hypothetical protein	50% (55/111)	*Paracoccus aminophilus*	WP_020950506
42	36362-38416	→	684	type III secretion system component SctV	69% (468/681)	*Paracoccus aminophilus*	WP_020950505

Gene no.	Coding region (bp)	Strand	Protein size (aa)	Possible function	Best BLAST hits		
					% identity (aa)	Organism	GenBank accession no.
43	38455-39675	→	406	type III secretion system component SctC2	64% (248/390)	*Paracoccus aminophilus*	WP_020950504
44	39672-40079	→	135	hypothetical protein	48% (45/94)	*Paracoccus* sp. 39524	KGJ16302
45	40174-41154	→	326	type III secretion system helper protein, lytic transglycosylase	54% (130/320)	*Advenella mimigardefordensis* DPN7	AHG62189
46	41322-45302	→	1326	hypothetical protein	59% (713/1214)	*Inquilinus limosus* MP06	KGM31571
47	45426-46037	→	203	type III secretion system component SctC1	42% (81/194)	*Paracoccus aminophilus*	WP_020950524
48	46030-46338	→	102	type III secretion system protein	55% (50/91)	*Paracoccus aminophilus*	WP_020950523
49	46386-46742	→	118	type III secretion system protein	39% (46/117)	*Paracoccus aminophilus*	WP_020950521
50	46755-47534	→	259	type III secretion system component SctJ	65% (149/228)	*Paracoccus aminophilus*	WP_020950520
51	47531-48217	→	228	type III secretion system protein	36% (43/121)	*Paracoccus aminophilus*	WP_020950519
52	48217-48819	→	200	type III secretion system component SctL	51% (94/184)	*Paracoccus aminophilus*	WP_020950518
53	48834-50138	→	434	type III secretion system component SctN	76% (318/419)	*Paracoccus aminophilus*	WP_020950517
54	50140-50634	→	164	type III secretion system protein	34% (48/142)	*Paracoccus aminophilus*	WP_020950516
55	50631-51821	→	396	type III secretion system component SctQ	41% (154/377)	*Paracoccus aminophilus*	WP_020950515
56	51821-52465	→	214	type III secretion system component SctR	76% (162/214)	*Paracoccus aminophilus*	WP_020950514
57	52472-52726	→	84	type III secretion system component SctS	80% (68/84)	*Paracoccus aminophilus*	WP_020950513
58	52723-53559	→	278	type III secretion system component SctT	52% (131/250)	*Paracoccus aminophilus*	WP_020950512
59	53556-54578	→	340	type III secretion system component SctU	51% (170/336)	*Paracoccus aminophilus*	WP_020950511
60	54645-55070	→	107	transposase OrfA (truncated copy of an IS)	99% (106/107)	*Paracoccus* sp. 39524	KGJ18543
61	55304-55987	→	227	transposase (IS*Pko6*c)	99% (226/227)	*Paracoccus* sp. TRP	WP_010398086
62	56240-57559	←	439	toxin HipA	100% (439/439)	*Paracoccus* sp. TRP	WP_010400321

Gene no.	Coding region (bp)	Strand	Protein size (aa)	Possible function	Best BLAST hits		
					% identity (aa)	Organism	GenBank accession no.
63	57559-57810	←	83	antitoxin HipB, transcriptional regulator XRE family	100% (83/83)	*Paracoccus* sp. TRP	WP_010400320
64	58290-59129	→	279	transposase (truncated copy of an IS)	88% (245/279)	*Sphingobium baderi* LL03	EQA98215
65	59432-60361	←	309	2'-hydroxyisoflavone reductase	100% (309/309)	*Paracoccus* sp. TRP	WP_029372744
66	60412-61815	←	467	aldehyde dehydrogenase	100% (467/467)	*Paracoccus* sp. TRP	WP_010400316
67	61867-62325	←	152	hypothetical protein	100% (139/139)	*Paracoccus* sp. TRP	WP_029372742
68	62499-63404	→	301	transcriptional regulator LysR family	100% (301/301)	*Paracoccus* sp. TRP	WP_010400311
69	63582-63998	←	138	truncated outer membrane protein-like	100% (138/138)	*Paracoccus* sp. TRP	WP_010400310
70	64304-64693	←	129	transcriptional regulator HxlR family	100% (129/129)	*Paracoccus* sp. TRP	WP_010400309
71	64810-65493	→	227	transposase (IS*Pko6*d)	99% (225/227)	*Paracoccus* sp. TRP	WP_010398086
72	65839-67206	←	455	alpha,alpha-trehalose-phosphate synthase OtsA	85% (386/453)	*Aquamicrobium defluvii*	EXL01463
73	67268-68098	←	273	trehalose-6-phosphate phosphatase OtsB	53% (119/224)	*Rhizobium leucaenae*	WP_028750974
74	68266-68892	←	208	truncated transposase	99% (189/191)	*Paracoccus pantotrophus*	WP_024845867
75	69475-69834	→	119	transposase OrfA (IS*Pko4*)	99% (118/119)	*Paracoccus* sp. TRP	WP_017998939
76	69857-70837	→	326	transposase OrfB (IS*Pko4*)	94% (306/326)	*Sphingobium japonicum*	WP_006948954
77	70868-71110	←	80	truncated transposase	87% (66/76)	*Paracoccus* sp. TRP	WP_010398086
78	71338-72366	←	342	transposase (IS*Pko7*)	100% (342/342)	*Paracoccus* sp. J55	WP_028714619
79	72875-75703	←	942	hypothetical protein	59% (550/931)	*Paracoccus aminophilus*	WP_020950528
80	75762-76181	→	139	truncated transposase	54% (97/180)	*Agrobacterium tumefaciens*	CDN95874
81	76282-76968	→	228	truncated transposase	89% (112/126)	*Dinoroseobacter shibae*	WP_012187069
82	76982-77665	←	227	transposase (IS*Pko6*)	99% (225/227)	*Paracoccus* sp. TRP	WP_010398086
83	78317-79096	→	259	hypothetical protein	99% (225/228)	*Thioclava dalianensis*	KEP67870
84	79259-79969	←	236	transposase (IS*Ppa9*a)	99% (234/236)	*Paracoccus* sp. N5	WP_026155209

Gene no.	Coding region (bp)	Strand	Protein size (aa)	Possible function	Best BLAST hits		
					% identity (aa)	Organism	GenBank accession no.
85	80262-81821	→	519	type I restriction system protein HsdM	92% (480/519)	*Pseudomonas aeruginosa*	WP_023127147
86	81818-83005	→	395	type I restriction system protein HsdS	63% (261/412)	*Klebsiella pneumoniae*	WP_032437415
87	83007-84209	→	400	anticodon nuclease PrrC	92% (366/400)	*Bordetella avium*	WP_012416240
88	84206-87304	→	1032	type I restriction system protein HsdR	94% (967/1032)	*Burkholderia multivorans*	AIO75074
89	87488-87796	←	102	truncated transposase	100% (102/102)	*Paracoccus* sp. TRP	WP_010398086
90	87808-88170	←	120	truncated transposase	97% (113/116)	*Paracoccus* sp. TRP	WP_010398086
91	88520-89452	→	310	transposase (IS*Pko1*b)	99% (309/310)	*Paracoccus* sp.	WP_028710025
92	89599-90090	→	163	nuclease	54% (76/142)	alpha proteobacterium BAL199	WP_007681224
93	90667-91215	←	182	transposase OrfB (IS*Pko5*e)	86% (148/172)	*Paracoccus pantotrophus*	ACG76127
94	91017-91442	←	141	transposase OrfA (IS*Pko5*e)	95% (134/141)	*Paracoccus pantotrophus*	WP_024844571
95	91576-92055	←	159	transcriptional regulator MarR family	77% (122/158)	*Aminobacter* sp. J41	WP_024846690
96	92101-93615	→	504	FAD-binding domain-containing monooxygenase	70% (353/501)	*Aminobacter* sp. J41	WP_024846691
97	93838-95049	→	403	major facilitator superfamily transporter	67% (271/403)	*Agrobacterium tumefaciens*	KAJ33607

Table S3.
Complete and partial ISs identified within pKON1.

Designation	Position	Orientation	Size (bp)	Identified IS			DR (sequence)	IS family/ IS group
				IRL	*tnp* gene(s)	IRR		
IS*Pme2*a	14-864	→	851	+	+	+	TA	IS*5* / IS*427*
IS*Pko2*	5741-7136	→	1396	+	+	+	GCGCCTTT	IS*256*
IS_trunc_1	8559-8666[b]	→	108[b]	-	partial	-	-	IS*66*
IS_trunc_2	8667-8924	→	258	+	partial	-	-	IS*3*/IS*3*
IS_trunc_3 (IS*Ppa9*')[a]	9631-9281	←	351	-	partial	+	-	IS*6*
IS*Pko6*	9632-10448	→	817	+	+	+	-	IS*6*
IS_trunc_4	11493-12588	→	1096	+	partial	-	-	IS*3*/IS*51*
IS*Pko5*	13439-12589	←	851	+	+	+	TA	IS*5*/IS*427*
IS_trunc_5 (IS*Pko6*')[a]	13458-14053	→	615	-	+	+	-	IS*6*
IS*Pko5*a	16690-17540	→	851	+	+	+	TA	IS*5*/IS*427*
IS_trunc_6	18668-18516	←	153	+	partial	+	-	IS*5*/IS*427*
IS*Pko6*a	18669-19485	→	817	+	+	+	-	IS*6*
IS*Pko3*	24122-25171	→	1050	+	+	+	-	IS*5*/IS*903*
IS*Pko5*	26602-27452	→	851	+	+	+	TA	IS*5*/IS*427*
IS*Pko4*	27915-29438	→	1524	+	+	+	-	IS*3*/IS*150*
IS*Pko6*b	31235-32051	→	817	+	+	+	-	IS*6*
IS*Pko5*b	33850-33001	←	850	+	+[c]	+	TA	IS*5*/IS*427*
IS_trunc_7 (IS*Pko5*')[a]	54578-55100	→	493	+	partial	-	-	IS5/IS427
IS_trunc_8 (IS*Pko6*')[a]	55223-55101	←	123	-	partial	+	-	IS*6*

Designation	Position	Orientation	Size (bp)	Identified IS			DR	IS family/
				IRL	*tnp* gene(s)	IRR	(sequence)	IS group
IS*Pko6*c	55224-56040	→	817	+	+	+	-	IS*6*
IS_trunc_9	58104-59129[b]	→	1026[b]	-	partial	-	-	IS*701*
IS*Pko6*d	64730-65546	→	817	+	+	+	-	IS*6*
IS_trunc_10	65547-65705[b]	→	159[b]	-	partial	-		IS*3*/IS*150*
IS_trunc_11 (IS*Pko1*')[a]	68841-68212	←	630	-	partial	+	-	IS*5*/IS*903*
IS_trunc_12 (IS*Pko6*')[a]	69388-68842	←	547	-	partial	+	-	IS*6*
IS*Pko4*a	69389-70912	→	1524	+	+	+	TTTCCTGG	IS3/IS150
IS_trunc_13 (IS*Pko6*')[a]	71190-70883	←	278	+	partial	-	-	IS*6*
IS*Pko7*	72409-71280	←	1130	+	+	+	-	IS*110*/ IS*1111*
IS_trunc_14	75762-76248[b]	→	487[b]	-	partial	-	-	IS*66*
IS_trunc_15	76249-76982	→	680	+	partial	-	-	IS*3*/IS*3*
IS*Pko6*	77745-76929	←	817	+	+	+	-	IS*6*
IS*Ppa9*a	80017-79206	←	812	+	+	+	-	IS*6*
IS_trunc_16 (IS*Pko5*')[a]	87304-87434	→	131	+	partial	-	-	IS*5*/IS*427*
IS*Pko6*e	88250-87434	←	816	+	+[c]	+	-	IS*6*
IS*Pko1*b	88457-89506	→	1103	+	+	+	-	IS*5*/IS*903*
IS_trunc_17	90546-90656	→	111	+	partial	-	-	IS5/IS903
IS*Pko5*c	91511-90657	←	851	+	+	+	TA	IS*5*/IS*427*

[a] truncated IS (complete IS copy within pKON1)
[b] position and length of partial *tnp* gene
[c] *tnp* gene with frameshift mutation

Publikacja 2

Łukasz Dziewit, **Jakub Czarnecki**, Daniel Wibberg, Monika Radlińska, Paulina Mrozek, Michał Szymczak, Andreas Schlüter, Alfred Pühler, Dariusz Bartosik

Architecture and functions of a multipartite genome of the methylotrophic bacterium *Paracoccus aminophilus* JCM 7686, containing primary and secondary chromids

BMC Genomics, 2014, 15:124

Impact Factor: 3,986
Punktacja MNiSW: 40

Dziewit et al. BMC Genomics 2014, 15:124
http://www.biomedcentral.com/1471-2164/15/124

BMC Genomics

RESEARCH ARTICLE Open Access

Architecture and functions of a multipartite genome of the methylotrophic bacterium *Paracoccus aminophilus* JCM 7686, containing primary and secondary chromids

Lukasz Dziewit[1*], Jakub Czarnecki[1], Daniel Wibberg[2], Monika Radlinska[3], Paulina Mrozek[3], Michal Szymczak[1], Andreas Schlüter[2], Alfred Pühler[2] and Dariusz Bartosik[1]

Abstract

Background: *Paracoccus aminophilus* JCM 7686 is a methylotrophic α-*Proteobacterium* capable of utilizing reduced one-carbon compounds as sole carbon and energy source for growth, including toxic *N,N*-dimethylformamide, formamide, methanol, and methylamines, which are widely used in the industry.
P. aminophilus JCM 7686, as many other *Paracoccus* spp., possesses a genome representing a multipartite structure, in which the genomic information is split between various replicons, including chromids, essential plasmid-like replicons, with properties of both chromosomes and plasmids. In this study, whole-genome sequencing and functional genomics approaches were applied to investigate *P. aminophilus* genome information.

Results: The *P. aminophilus* JCM 7686 genome has a multipartite structure, composed of a single circular chromosome and eight additional replicons ranging in size between 5.6 and 438.1 kb. Functional analyses revealed that two of the replicons, pAMI5 and pAMI6, are essential for host viability, therefore they should be considered as chromids. Both replicons carry housekeeping genes, e.g. responsible for *de novo* NAD biosynthesis and ammonium transport. Other mobile genetic elements have also been identified, including 20 insertion sequences, 4 transposons and 10 prophage regions, one of which represents a novel, functional serine recombinase-encoding bacteriophage, ϕPam-6. Moreover, *in silico* analyses allowed us to predict the transcription regulatory network of the JCM 7686 strain, as well as components of the stress response, recombination, repair and methylation machineries. Finally, comparative genomic analyses revealed that *P. aminophilus* JCM 7686 has a relatively distant relationship to other representatives of the genus *Paracoccus*.

Conclusions: *P. aminophilus* genome exploration provided insights into the overall structure and functions of the genome, with a special focus on the chromids. Based on the obtained results we propose the classification of bacterial chromids into two types: "primary" chromids, which are indispensable for host viability and "secondary" chromids, which are essential, but only under some environmental conditions and which were probably formed quite recently in the course of evolution. Detailed genome investigation and its functional analysis, makes *P. aminophilus* JCM 7686 a suitable reference strain for the genus *Paracoccus*. Moreover, this study has increased knowledge on overall genome structure and composition of members within the class *Alphaproteobacteria*.

Keywords: *Paracoccus aminophilus* JCM 7686, Genome, Chromid, Plasmid, Mobile genetic element, Bacteriophage

* Correspondence: ldziewit@biol.uw.edu.pl
[1]Department of Bacterial Genetics, Institute of Microbiology, Faculty of Biology, University of Warsaw, Miecznikowa 1, 02-096 Warsaw, Poland
Full list of author information is available at the end of the article

Background

The genus *Paracoccus* (*Alphaproteobacteria*) currently comprises 40 recognized and validly named species, isolated from different environments in various geographical locations. Members of this genus exhibit a broad range of metabolic flexibility, especially in respiratory processes, e.g. employing nitrate, nitrite, nitrous oxide and nitric oxide as alternative electron acceptors in denitrification and the ability to use one-carbon (C1) compounds (e.g. methanol, methylamine) as electron donors to respiratory chains [1]. Moreover, *Paracoccus* spp. as facultative chemolithoautotrophs may utilize reduced sulfur compounds (e.g. thiocyanate, thiosulfate or elemental sulfur), molecular hydrogen and Fe(II) as energy sources [2-4]. *Paracoccus* spp. are also able to use a broad range of organic compounds as their sole source of carbon and energy, including pollutants such as acetone, dichloromethane, formamide, *N,N*-dimethylformamide (DMF) and methylamine [3,5]. Another common feature of *Paracoccus* spp. is methylotrophy, defined as the ability to utilize reduced C1 carbon substrates containing no carbon-carbon bonds (including methane, methanol, methylated amines, halogenated methanes and methylated sulfur species) as their sole source of carbon and energy for growth.

Having a versatile metabolism, *Paracoccus* spp. play an important role in biogeochemical cycles and they have also been successfully employed in the biotreatment of contaminated environments, e.g. bioremediation of soils contaminated with polycyclic aromatic hydrocarbons (PAHs) using *Paracoccus* sp. HPD-2 [6].

Although *Paracoccus* spp. constitute an interesting and metabolically versatile group of bacteria with substantial biotechnological potential, little is known about the content and organization of their genomes. Only one complete genome of *Paracoccus denitrificans* PD1222 has been deposited in the NCBI database ([GenBank:CP000489], [GenBank: CP000490] and [GenBank:CP000491]). This genome is composed of two chromosomes (ChI – 2.9 Mb and ChII – 1.7 Mb) and a single megaplasmid (plasmid 1) of 653 kb. Four *Paracoccus* spp. genome sequencing projects (*Paracoccus* sp. TRP [7], *Paracoccus denitrificans* SD1 [8], *Paracoccus* sp. N5 and *P. zeaxanthinifaciens* ATCC 21588) are currently in progress.

Much more is known about mobile genetic elements (MGEs) of *Paracoccus* spp. Baj and colleagues (2000) [9] demonstrated that bacteria belonging to this genus usually harbor at least one plasmid (in most cases a mega-sized replicon, exceeding 100 kb). Several multireplicon strains carrying 4 or more plasmids were identified, including *P. aminophilus* JCM 7686. Our group has already obtained the sequences of 18 plasmids, ranging in size from 2.7 to 40 kb, that originate from different *Paracoccus* spp.: *P. methylutens* DM12 (2 plasmids) [10], *P. pantotrophus* DSM 11072 (1 plasmid) [11] and four carotenoid producers, *P. aestuarii* DSM 19484 (5 plasmids) ([GenBank:JQ041633], [GenBank:JQ065021], [GenBank:JQ066766], [GenBank:JQ684025], [GenBank: JQ796370]), *P. haeundaensis* LMG P-21903 (2 plasmids) ([GenBank:JQ066767], [GenBank:JQ684024]), *P. marcusii* DSM 11574 (5 plasmids) ([GenBank: KC542384], [GenBank: KC561053], [GenBank:KC561054], [GenBank:KC561055], [GenBank:JQ796371) and *P. marcusii* OS22 (3 plasmids) ([GenBank:JQ664550], [GenBank:JQ678602], [GenBank: JQ684023]). We have also sequenced the basic replicons of three other plasmids from *P. alcaliphilus* JCM 7364 (pALC1) and *P. versutus* UW1 (pTAV1 and pTAV3) [12-14]. Moreover, complex analyses of 25 *Paracoccus* spp. strains using trap plasmid systems have led to the identification and characterization of (i) 48 insertion sequences (ISs), (ii) a composite transposon Tn*6097* carrying genetic modules involved in the arginine deiminase pathway and daunorubicin/doxorubicin resistance, (iii) 3 non-composite transposons of the Tn*3* family, (iv) a transposable genomic island Tn*Ppa1* (45 kb) and (v) several transposable modules (TMos) generated by a single copy of the IS*1380* family insertion sequence [10,15-20]. The findings outlined above suggest that horizontal gene transfer (HGT) events occur frequently in *Paracoccus* spp. genomes, which may explain their metabolic flexibility.

Among *Paracoccus* species, *P. aminophilus* JCM 7686 is of particular interest since it is a methylotrophic bacterium capable of utilizing several toxic C1 compounds, including *N,N*-dimethylformamide, formamide as well as tri-, di- and monomethylamine, which are widely used in the chemical industry [21]. *P. aminophilus* JCM 7686 carries eight indigenous, extrachromosomal replicons (pAMI1 to pAMI8) ranging in size from 5.6 kb to approximately 440 kb. Our previous analyses, focused on the three smallest plasmids, pAMI3 (5.6 kb), pAMI2 (18.6 kb) and pAMI7 (20.5 kb), revealed the presence of (i) novel types of plasmid-encoded maintenance systems [22,23], (ii) a type II restriction-modification module with NcoI specificity [24] and (iii) genes crucial for the first step in the degradation of DMF [25].

In the present study, the genome of *P. aminophilus* JCM 7686 was completely sequenced and analyzed. An in-depth exploration of this genome sequence, followed by functional analyses, provided considerable insights into its overall architecture, as well as the functions of particular replicons. Moreover, in this study, the first inducible *Paracoccus* phage was identified.

Results and discussion

Sequencing and general features of the *P. aminophilus* JCM 7686 genome

A 454-pyrosequencing run for the *P. aminophilus* JCM7686 genomic DNA yielded 598348 shotgun and

8-kb-long paired-end reads with a total number of 225,950,536 bp that were assembled into 17 scaffolds. The scaffolds consisted of 429 large (> 500 nucleotides) and 129 small (100–500 nucleotides) contigs. The gaps in the chromosome and plasmids were closed by a PCR-based approach followed by sequencing of the corresponding amplicons. Subsequently, the genome was re-sequenced (for the quality check) applying the Illumina HiScanSQ Genome Analyzer.

The genome of *P. aminophilus* JCM 7686 is composed of a single circular chromosome of 3,613,807 bp and eight circular plasmids: pAMI1 (118,164 bp), pAMI2 (18,563 bp), pAMI3 (5575 bp), pAMI4 (438,126 bp), pAMI5 (294,017 bp), pAMI6 (206,583 bp), pAMI7 (20,542 bp) and pAMI8 (202,421 bp) (Figure 1). Thus, the total size of the genome is 4,917,798 bp. The overall GC content of the chromosome is 63.4%, which is consistent with other sequenced *Paracoccus* genomes. The GC contents of the JCM 7686 plasmids ranges between 57.4% (pAMI7) and 64.2% (pAMI4) (Table 1).

The genome of JCM 7686 possesses 59 tRNAs and four clusters of 5S, 16S and 23S rRNA genes, which are located within the chromosome and plasmid pAMI1 (Table 1). This strain can produce tRNAs for all 20 amino acids, and genes encoding all the aminoacyl-tRNA synthetases are present within its chromosome. Moreover, a single tRNA for selenocyteine (tRNA-SeC) was also identified.

The JCM 7686 genome contains 4573 putative coding sequences (CDSs): 3416 in the chromosome and 1157 in the plasmids (Table 1). The coding density within the chromosome is 90.2%, while within the plasmids it varies from 72% (pAMI3) to 92.8% (pAMI6) (Table 1). We could assign putative biological functions to 3506 CDSs (76.6%), while 1069 CDSs (23.4%) were annotated as encoding hypothetical proteins of unknown function.

Predicted *P. aminophilus* proteins were functionally categorized and the proportions in each COG category were calculated. A total of 1550 (33.9%) predicted proteins were described as involved in the overall cellular metabolism, being assigned to COG functional categories C, G, E, F, H, I or Q. It appeared that more than 40% of the genes encoding proteins involved in metabolic processes are located within extrachromosomal elements,

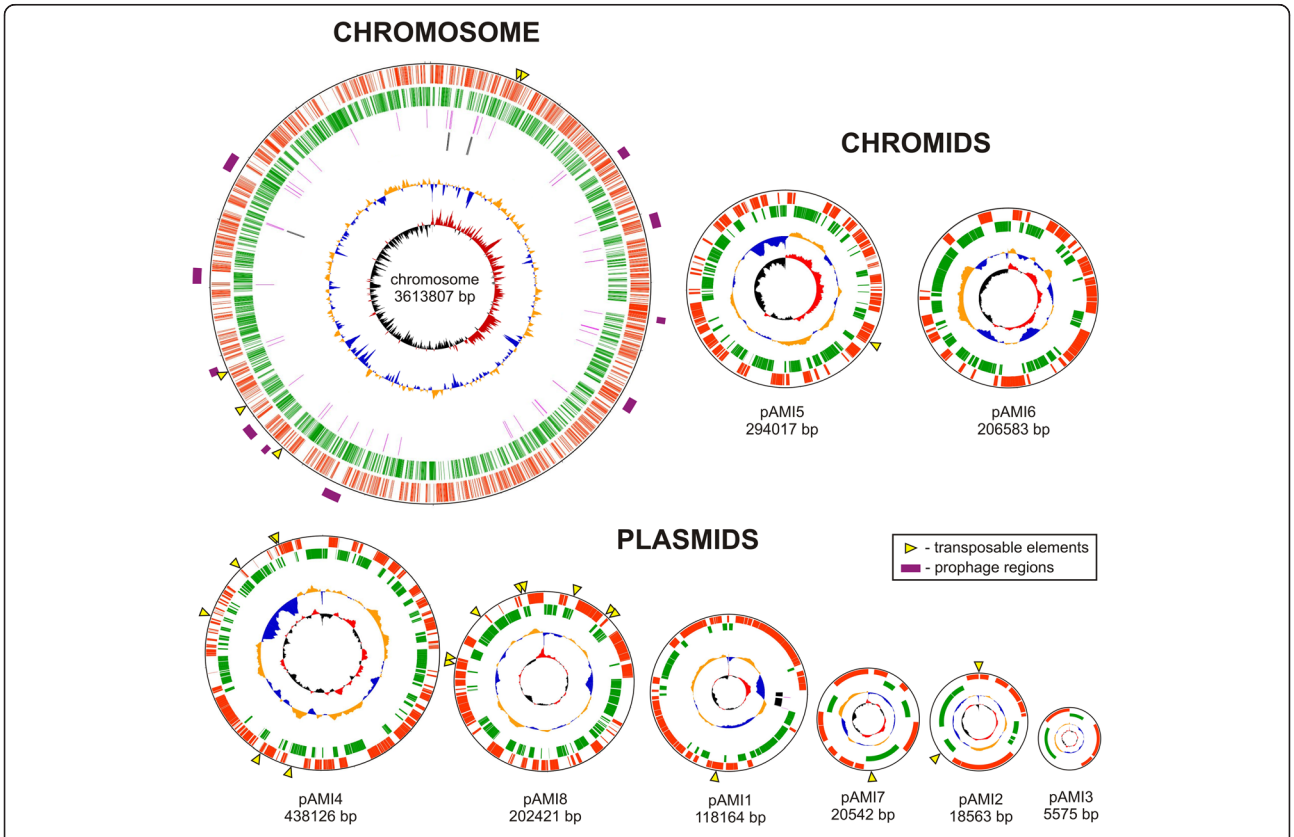


Figure 1 Circular representations of the *P. aminophilus* JCM 7686 genome. Circles displayed (from the outside): (i) predicted CDSs transcribed in the clockwise direction, (ii) predicted CDSs transcribed in the counterclockwise direction, (iii) the position of tRNA genes, (iv) the position of rRNA genes, (v) the GC percent deviation, (vi) GC skew (G + C/G-C). The tRNA and rRNA genes are present only within the chromosome and plasmid pAMI1. Yellow arrow heads indicate transposable elements. Violet lines represents the prophage-like regions. Circles are drawn not in scale.

Table 1 General features of the *P. aminophilus* JCM 7686 genome

General features	Chromosome	pAMI1	pAMI2	pAMI3	pAMI4	pAMI5	pAMI6	pAMI7	pAMI8
Size (bp)	3,613,807	118,164	18,563	5575	438,126	294,017	206,583	20,542	202,421
GC content (%)	63.4	63.3	62	60.8	64.2	62.8	63.9	57.4	62.3
Coding density (%)	90.2	87.5	79.1	72	88.1	88.5	92.8	83.6	89.2
Number of ORFs	3416	99	19	7	375	264	171	19	203
Number of tRNA genes	56	3	0	0	0	0	0	0	0
Number of 16S-23S-5S rRNA operons	3	1	0	0	0	0	0	0	0
Phage regions	9	0	0	0	0	0	0	0	0
GTA regions	1	0	0	0	0	0	0	0	0
Transposases (including truncated)	25	1	4	0	20	4	0	1	17
Complete non-composite transposons	0	1	0	0	0	0	0	1	2
Complete insertion sequences	5	0	2	0	6	1	0	0	6

mainly pAMI1, pAMI4, pAMI5 and pAMI6 (426 genes in total). This finding suggests that these four replicons perform an important role (Additional file 1).

The most abundant class of proteins comprises those involved in amino acid transport and metabolism (COG category E). A total of 576 (12.6%) CDSs were assigned to this category, of which 191 occur within the aforementioned replicons: pAMI1 (41 genes), pAMI4 (65), pAMI5 (52) and pAMI6 (31) (Additional file 1).

Transcription regulatory network

Within the JCM 7686 genome, we identified 12 genes (7 in the chromosome, 3 in pAMI4, 1 in pAMI1 and 1 in pAMI5) encoding predicted Sigma factors: (i) FecI-like σ^{19} – required for the uptake of iron (5 copies, JCM7686_1690, JCM7686_1812, JCM7686_pAMI1p074, JCM7686_pAMI4p059 and JCM7686_pAMI5p016), (ii) RpoE-like σ^{24} – enabling the expression of genes involved in the heat shock response, as well as extracytoplasmic transport (4 copies, JCM7686_2748, JCM7686_2865, JCM7686_3466 and JCM7686_pAMI4p292), (iii) RpoH-like σ^{32} – involved in the heat shock and other stress responses (JCM7686_2477), (iv) RpoN-like σ^{54} – involved in nitrogen metabolism (JCM7686_pAMI4p216) and (v) RpoD-like σ^{70} – regulating gene expression during exponential growth (JCM7686_2204). Genes equivalent to *rpoF* (σ^{28}) and *rpoS* (σ^{38}), responsible for the expression of flagellar genes and genes in the stationary phase of growth, respectively, were not detected in this genome. In addition, we identified anti-sigma (JCM7686_3151) and anti-anti-sigma (JCM7686_3150) factors within the JCM 7686 chromosome.

The presence of eleven alternative sigma factors (of four types – σ^{19}, σ^{24}, σ^{32} and σ^{54}), enables global changes in gene expression and permits alteration of the overall cellular metabolism in response to changing environmental conditions [26]. Thus, this feature indirectly reflects the metabolic flexibility and adaptive abilities of the strain JCM 7686. It is noteworthy that three of five *P. aminophilus fecI* (σ^{19}) genes were identified within plasmids, which suggests an important role for these replicons in iron transport and metabolism. Moreover, the only *rpoN* gene copy was carried by a plasmid, pAMI4. It was previously demonstrated that RpoN, besides regulating genes involved in nitrogen metabolism, enables the transcription of a wide range of other genes encoding proteins participating in the regulation of virulence-related factors, as well as in amino acid, carbohydrate and organic acid synthesis, utilization and transport [26]. This indicates that pAMI4 may be highly significant in overall gene regulation in the strain JCM 7686.

Global regulators, which have the ability to regulate operons that belong to various metabolic pathways, are another basic component of transcription regulatory networks in bacteria. It has been shown that the expression of 51% of the *E. coli* genes is under the control of only seven regulatory proteins: CRP, FNR, IHF, FIS, ArcA, NarL and Lrp [27]. Within the JCM 7686 genome we identified genes encoding 5 types of putative global gene regulators: (i) three CRP/FNR family transcriptional regulators (JCM7686_1106, JCM7686_3086, JCM7686_pAMI8p144), (ii) three FIS family transcriptional regulators (JCM7686_3395, JCM7686_pAMI4p202, JCM7686_pAMI5p202), (iii) a H-NS proteins (JCM7686_3426, JCM7686_pAMI8p115), (iv) a Lrp regulator (JCM7686_0212) and (v) integration host factor (IHF) subunits A (IhfA; JCM7686_1093) and B (IhfB; JCM 7686_1288).

We also identified 217 genes encoding transcriptional regulators with predicted local specificity of action, probably limited to a single gene/gene cluster (Additional file 2). A quarter of the identified regulators were assigned to the LysR family – the most abundant type of transcriptional regulators in bacteria [28]. Interestingly, we identified genes encoding 8 regulators of the LuxR family, which

are located within the chromosome and plasmids pAMI1, pAMI2, pAMI4 and pAMI8. The majority of LuxR-type proteins represents transcription activators, which specifically bind to *N*-acyl homoserine lactones (AHL; synthesized by a LuxI protein) that are secreted signaling molecules involved in quorum sensing in a variety of Gram-negative bacteria (e.g. [29]). We also identified CDSs for 3 putative *N*-acyl-*L*-homoserine lactone synthetases (LuxI-like proteins): two localized within the chromosome (JCM7686_2124, JCM7686_3181) and one in plasmid pAMI1 (JCM7686_pAMI1p026). The presence of *luxR* and *luxI* genes strongly suggests that quorum sensing plays a role in regulation of *P. aminophilus* gene expression.

Two-component systems are another important part of bacterial transcription regulatory networks present in *P. aminophilus*. Such systems are directly involved in sensing a cell's external environment and signal transduction (e.g. [30]). Within the JCM 7686 genome we identified 23 pairs of genes encoding histidine protein kinases (HPKs) and phospho-aspartyl response regulators: 15 of them within the chromosome and others located on plasmids pAMI4, pAMI5, pAMI6 and pAMI8 (Additional file 3). Moreover, we identified 6 HPKs with unknown partner response regulators in the chromosome and pAMI6 (Additional file 3).

Among the *P. aminophilus* HPKs we found homologs of enzymes of well-described two-component phosphorelay systems involved in regulating (i) nitrogen assimilation – NtrB (JCM7686_0575 and NtrX (JCM7686_0577) [31], (ii) chemotaxis – CheA (JCM7686_1281) [32], (iii) phosphate homeostasis – PhoR (JCM7686_2063) [33], (iv) differentiation and cell cycle progression – CckA (JCM7686_2539) [34], (v) C4-dicarboxylate metabolism – DctB (JCM7686_2824) [35], (vi) expression of virulence factors – QseC (JCM7686_3369; JCM7686_pAMI5p117) [36], (vii) methanol and formaldehyde oxidation – FlhS (JCM7686_3383), (viii) several anaerobic processes and assimilation of CO_2 and N_2 – RegB (JCM7686_3423) [37], (ix) high affinity potassium-uptake – KdpD (JCM7686_pAMI4p337) [38] and (x) expression of the *tor* structural operon encoding the trimethylamine *N*-oxide reductase respiratory system in response to substrate availability – TorS (JCM7686_pAMI5p018) [39].

Stress response

In the majority of cases, the processes underlying the global stress response of bacteria are dependent on alterations in gene expression, usually controlled at the transcriptional level by various sigma factors [40]. However, there are also several more specific stress response mechanisms.

A fundamental trigger of the cellular stress response is DNA damage. In this scenario, the bacterial defense mechanism relies on a conserved inducible pathway – the SOS response (e.g. [41]). The primary components of this pathway, RecA recombinase and the LexA repressor, are encoded within the JCM 7686 chromosome (JCM7686_2538 and JCM7686_0753, respectively).

Moreover, the SpoT/RelA, (p)ppGpp synthetase I (JCM7686_1948), involved in the bacterial stringent response (triggered by nutritional deprivation) was also identified. During the stringent response process, an accumulation of the signaling nucleotides pppGpp and ppGpp occurs, which causes dramatic alterations in gene expression [42]. An important role in this process is also played by the chromosomally-encoded RelE toxin of the *relBE* systems, which, as mRNA-cleaving enzymes, globally inhibits translation during amino acid starvation [43]. Within the *P. aminophilus* genome we identified 11 toxin-antitoxin (TA) pairs representing the families *relBE/parDE* (5 TA systems), *phd-doc* (1), *ccdAB* (1), *hipAB* (2), and 2 hybrid systems (*vapC-phd*). The majority of the identified TA modules were localized within the plasmids (Additional file 4), and they most probably perform replicon stabilization functions. Only a single TA system of the *hipAB* family was located on the chromosome.

Among the most important factors in the maintenance of cellular fitness under changing environmental conditions are molecular chaperones. They are involved in various processes in bacterial cells, such as assisting the folding of newly synthesized proteins, protein secretion, preventing the aggregation of proteins under heat shock conditions, and repairing proteins that have been damaged or misfolded due to stress conditions (e.g. [44]). Within the *P. aminophilus* genome we identified 35 putative molecular chaperones, the majority of which (32) are encoded by the chromosome. These include heat and cold shock proteins, proteases, disulfide bond chaperones and protein-export chaperones (Additional file 5). We also identified a major RNA chaperone, Hfq, which is a key player in small RNA (sRNA)-mediated regulation of target mRNAs, that promotes sRNA-mRNA base pairing to permit rapid adaptive responses [45]. Surprisingly, we did not identify homologs of the bacterial HtpG (Hsp90) protein, which is responsible for heat or chemical shock responses. This is unusual because genes encoding HtpG proteins are present in the majority of bacterial genomes [46].

DNA recombination and repair

Since the integrity of the genomic DNA is fundamental to bacterial persistence, it is extremely important that the cell is equipped with sufficient DNA repair proteins to protect it against genetic damage [47]. Within the JCM 7686 genome we identified 59 genes encoding proteins predicted to be involved in various DNA repair pathways. The vast majority of them (47) are located within the

chromosome. We performed a detailed characterization and classification of all genes encoding DNA repair-related proteins (summarized in Additional file 6). Comparison of the DNA repair genes of *P. aminophilus* with those of two well defined species, *Escherichia coli* (*Gammaproteobacteria*) and *Caulobacter crescentus* (*Alphaproteobacteria*), revealed the presence of 53 and 58 orthologous genes, respectively (Additional file 6).

As in *C. crescentus* [47], the JCM 7686 genome contains no *mutH* and *dam* homologs, which implies that both strains use different proteins to recognize and repair DNA replication errors. *P. aminophilus*, like many other *Alphaproteobacteria* (including *C. crescentus*) also lacks the RecBCD module, but possesses the related system AddAB instead, which recognizes a 5-base Chi site (5′-AGCGG-3′; 17,218 such recognition sequences were identified within the JCM 7686 genome) [48]. Curiously, we also found that *P. aminophilus* does not encode a typical deoxyribodipyrimidine photolyase (Phr-like). These enzymes employ visible light as the energy source to monomerize pyrimidine dimers induced by UV irradiation [49], and orthologs are found within the *C. crescentus* genome. Instead of Phr, JCM 7686 encodes a non-related SplB-like DNA repair photolyase. It was shown that the SplB protein is a DNA repair enzyme responsible for the process of reversion of the thymine dimer, 5-thyminyl-5,6-dihydrothymine (spore photoproduct), formed during UV irradiation of *Bacillus subtilis* spores [50]. We hypothesize that this protein may function as the major photolyase of *P. aminophilus*, responsible for the photo-reactivation process.

Although the majority of the DNA repair proteins are found within the JCM 7686 chromosome, 12 genes encoding such proteins are present in three plasmids: pAMI4 (3 genes), pAMI5 (3) and pAMI8 (6). Among them is an *alkB* gene (carried by pAMI5), which encodes a highly conserved and usually chromosomally-encoded protein (AlkB), responsible for the repair of alkylation damage in DNA *via* an oxidative demethylation pathway [51]. The presence of this *alkB* gene within pAMI5 suggests that it may be an essential replicon.

DNA methylation

Genes encoding predicted DNA methyltransferases (MTases), which may play important roles in the regulation of replication, gene expression and mismatch repair systems (e.g. [52]) were detected in the JCM 7686 genome. In addition to the previously described M.PamI [24], that is part of the type II restriction-modification system encoded by plasmid pAMI7, we identified seven chromosomally-encoded putative MTase genes. All of them seem to be orphan MTases, since they lack associated partner endonucleases.

The gene JCM7685_3079 encodes a CcrM (cell-cycle regulating MTase) homolog. Its ability to modify adenine residues in GANTC sequences was confirmed *in vivo* and *in vitro* (Additional file 7). The CcrM methyltransferases were shown to be essential for the viability of various *Alphaproteobacteria* (including *Caulobacter crescentus*, *Sinorhizobium meliloti*, *Agrobacterium tumefaciens* and *Brucella abortus*) and to play a crucial role in the regulation of bacterial cell division (e.g. [53]). To confirm the pivotal role of the JCM7686_3079 gene product in the *P. aminophilus* cell cycle, we attempted to disrupt this gene. However, this proved to be impossible, unless a wild-type copy was provided *in trans*, demonstrating that this CcrM homolog is essential for the viability of the host strain.

The other six genes encoding putative MTases are located within predicted prophage regions. Based on automatic methyltransferase prediction algorithms we allocated the following genes to different MTase classes: m^6A MTases – JCM7686_1231, JCM7686_2255 and JCM7686_2934; m^4C MTase – JCM7686_0815; m^5C MTases – JCM7686_0772 and JCM7686_2655. Each of these predicted methylase genes was cloned into an expression vector and expressed in *E. coli*. The activity of the recombinant MTases was assessed and their sequence specificity determined using the endonuclease protection assay (according to [54]) (Additional files 8, 9, 10).

This analysis revealed that both m^5C MTases have relaxed substrate specificity [partial protection of sequences CCNGG, CCWGG, AGCT, CCGG, GCN_7GC and others (Additional file 8)], but they do not protect DNA against cleavage by R.PamI, encoded by the restriction-modification system of pAMI7. The m^4C MTase of JCM 7686 recognizes and methylates YGGCCR sequences (Additional file 9). The three m^6A MTases are highly similar (at least 95% aa sequence identity) and they methylate the sequence GANTC (Additional file 10). This is the same sequence specificity as assigned to the CcrM methylase (the main cell cycle regulator).

Multireplicon structure of the JCM 7686 genome

As mentioned above, the genome of *P. aminophilus* JCM 7686 is composed of a single chromosome and eight smaller replicons (pAMI1-pAMI8) (Figure 1). Predicted plasmids constitute 26.5% of the genome and they carry 1158 CDSs (about 25% of all JCM 7686 CDSs), which means that a huge amount of genetic information is stored within these replicons. Among the JCM 7686 plasmids we distinguished 5 mega-sized replicons (118–438 kb), pAMI1, pAMI4, pAMI5, pAMI6 and pAMI8, all carrying genes conserved in bacterial chromosomes, including chromosomes I and II of *P. denitrificans* PD1222 (Additional file 11). Since size is not an

infallible criterion for distinguishing plasmids from secondary chromosomes, functional analyses of the JCM 7686 replicons was undertaken.

To characterize the function of the mega-sized replicons in the cellular metabolism we constructed mini-derivatives, which were used to remove the native plasmids from the host cell by incompatibility (such analyses of the smallest plasmids pAMI2, pAMI3 and pAMI7 were performed previously [22,23,25]. The mini-derivatives were constructed by cloning DNA fragments containing the plasmid maintenance modules (including plasmid replication initiation modules), into the *E. coli*-specific, mobilizable, narrow-host-range vector pABW1 (ColE1-type *ori* of pMB1) [55]. Using the obtained mini-derivatives containing incompatibility determinants, we were able to remove plasmids pAMI1, pAMI4, pAMI6 and pAMI8 from JCM 7686 (Additional file 12). Interestingly, curing this strain of pAMI1 resulted in a change in colony morphology, which became irregular, flat and dry. We were unable to introduce the mini-derivative of pAMI5 into JCM 7686 cells. It was therefore impossible to remove pAMI5, which strongly suggests that this replicon might be essential to host viability.

Growth rate analysis of strains deprived of the individual megaplasmids was performed in rich LB medium, as well as in AC minimal salt medium. This revealed that the removal of plasmids pAMI1, pAMI4 and pAMI8 had no effect on the growth rate (Figure 2). Interestingly, the pAMI1-less derivative grew in the form of irregular flocks (Additional file 13), which were in fact the aggregations of bacterial cells. In the case of the pAMI6-less strain, growth in LB medium was significantly decreased, and it was completely inhibited in AC medium (Figure 2). This indicates that the presence of plasmid pAMI6 enables growth of this *P. aminophilus* strain in minimal medium.

Based on the obtained results we have classified the pAMI-replicons into two groups: (i) essential (pAMI5, pAMI6) and (ii) non-essential (pAMI1, pAMI2, pAMI3, pAMI4, pAMI7 and pAMI8) genetic elements.

Additionally, to extend the functional analyses, the host ranges of pAMI1-pAMI8 replicons were examined using their plasmid mini-derivatives cloned within plasmid pABW1. The tested strains of *Alphaproteobacteria* were classified into two orders: (i) *Rhodobacterales*, represented by *Paracoccus versutus* UW225 and *Paracoccus pantotrophus* KL100, and (ii) *Rhizobiales*, represented by the families *Brucellaceae* (*Ochrobactrum* sp. LM19R) and *Rhizobiaceae*, including members of the *Rhizobium*/*Agrobacterium* group (*Rhizobium etli* CE3 and *Agrobacterium tumefaciens* LBA288) and the *Sinorhizobium*/*Ensifer* group (*Sinorhizobium* sp. LM21R). Replication abilities of pAMI1-pAMI8 were also tested in *Alcaligenes* sp. LM16R (*Betaproteobacteria*) and *E. coli* BR825 (*Gammaproteobacteria*).

This analysis revealed that mini-derivatives of pAMI1, pAMI2, pAMI3, pAMI4, pAMI7 and pAMI8 could replicate in all analyzed strains of *Alphaproteobacteria*, while the introduction of pAMI5 and pAMI6 into *Sinorhizobium* sp. and *Ochrobactrum* sp. was impossible, which may be due to strong incompatibility with their native replicons. We were also unable to introduce any of the analyzed plasmids into representatives of the *Beta-* or *Gammaproteobacteria* (Additional file 14).

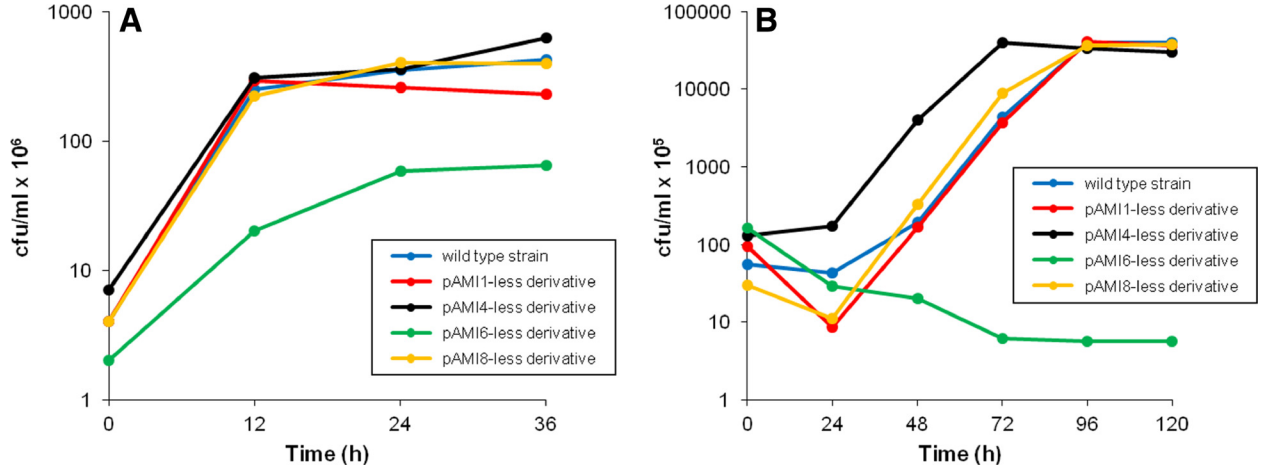


Figure 2 The growth rates of pAMI1-, pAMI4-, pAMI6- and pAMI8-less derivatives of *P. aminophilus* JCM 7686. Growth rates of the bacterial cultures were determined by viable cell counts (cfu).The strains [wild-type strain, pAMI1-less derivative, pAMI4-less derivative, pAMI6-less derivative and pAMI8-less derivative] were grown in 20 ml of liquid LB **(A)** or AC medium **(B)** with arabinose added as the sole source of carbon and energy. Samples were taken every 12 h for 36 h (in the case of LB medium) and every 24 h for 120 h (in the case of AC medium). The values plotted are the means of three replicates.

These findings indicate the relatively narrow host range of the analyzed replicons, being limited to members of the *Alphaproteobacteria* class.

Essential genetic elements – chromids

The curing experiments showed that the pAMI5 replicon could not be removed from JCM 7686 cells. Close inspection of the genetic content of this plasmid revealed that 62 (23.5%) of its genes are conserved in the chromosomes of *P. denitrificans* PD1222 (Additional file 11), and that 45 of them are singletons in the JCM 7686 genome. These genes encode several proteins of unknown function, but also proteins that are probably involved in the host metabolism, including (i) iron transport, (ii) D-pantothenate synthesis, (iii) motility, (iv) folate one-carbon metabolism and (v) *de novo* NAD biosynthesis. Interestingly, the *nadABC* (JCM7686_pAMI5p135-137) genes are the only ones encoding components of the *de novo* NAD biosynthetic pathway (Additional file 15) identified within the JCM 7686 genome, which may explain the essential nature of the pAMI5 replicon. It is important to mention that the loss of *nadABC* genes most probably cannot be compensated *via* the NAD salvage pathway, since the predicted NAD pyrophosphatase (an enzyme involved in the process) is also encoded by pAMI5 replicon (JCM7686_pAMI5p033 gene) (the remaining components of this pathway were identified in the JCM 7686 chromosome). Moreover, pAMI5 contains the only copies of (i) the *alkB* gene (JCM7686_pAMI5p075) induced during adaptive responses and involved in the direct reversal of alkylation damage [56] and (ii) the *hslJ* gene (JCM7686_pAMI5p111) encoding a heat shock protein [57].

All the aforementioned genes may play an important role in the JCM 7686 metabolism and we speculate that some of them may be responsible for the pAMI5 essentiality. However this needs to be experimentally confirmed.

Essential element status can also be assigned to the pAMI6 replicon. Although, we were able to remove it from JCM 7686 cells cultivated in LB medium, the growth of the pAMI6-less strain was completely inhibited in minimal salt medium (Figure 2). Analysis of the genetic content of pAMI6 revealed the presence of several modules that are typically encoded on chromosomes. Corresponding genes are involved in histidine, folate, glycerophospholipid, purine, selenoamino acid, sulfur, propanoate, glyoxylate and dicarboxylate metabolism. A surprising finding was the *amtB* gene (JCM7686_pAMI6p067, the only copy of this gene in the JCM 7686 genome), which encodes a NH_4^+ transporter of the Amt/Rh family [58]. Presence of this gene on pAMI6 may explain why this plasmid is essential for growth in minimal salt medium, where the only nitrogen source is NH_4Cl, which cannot be transported into the cell in the absence of an appropriate transporter. This finding was confirmed by the analysis of the strain growth in minimal medium with casamino acids as a nitrogen source instead of ammonium chloride. pAMI6-less strain was able to growth in such medium, but its growth rate was significantly decreased. This suggests that there are also some additional (to *amtB*) genes, which are responsible for the essentiality of this replicon.

A growing number of studies in the field of genomics have produced data suggesting that the structure of many bacterial genomes is more complex than previously assumed. Many bacteria bear additional, large, autonomous replicons, which (like chromosomes) carry a pool of housekeeping genes [59-62]. Unlike typical plasmids, such replicons are necessary for the viability of their hosts, and for this reason they were initially defined as "secondary chromosomes". However, this name is inadequate since these replicons possess many characteristics typical of plasmids: they contain plasmid-like replication systems and other genetic modules of plasmid origin. Bioinformatic analyses of these replicon sequences indicated that they were generated by the transfer of genetic information from chromosomes to plasmids co-residing in the cell. Due to their dualistic properties, they have been reclassified into a separate, newly distinguished group with properties of both chromosomes and plasmids: the chromids [60]. Our analyses indicate that pAMI5 and pAMI6 can be classified into this group of elements as well.

Non-essential genetic elements – plasmids

Six of the predicted JCM 7686 plasmids (including the previously described plasmids pAMI2, pAMI3 and pAMI7 [23,25]) were readily removed from the host cells. Among these dispensable replicons are the two *repABC*-type megaplasmids, pAMI4 and pAMI8 (Figure 1). Interestingly, pAMI4 and pAMI8 carry almost as many putative transposase genes (20 and 17, respectively; complete and truncated) as the whole chromosome (25 genes), and many more than the other replicons (0–4 genes) (Table 1). Within pAMI4 and pAMI8 we identified 101 genes encoding predicted transporters involved in the transport of (i) amino acids/dipeptides (*dpp*-like gene clusters), (ii) sn-glycerol-3-phosphate (*ugp*-like gene clusters), (iii) taurine (*tau*-like gene cluster) and (iv) inorganic ions [mainly iron (*fec*-like gene clusters), but also potassium, sodium and various heavy metals]. Genes encoding predicted transporters constitute 21 and 10% of the genetic information carried by plasmids pAMI4 and pAMI8, respectively. Moreover, within pAMI8 we identified a *vir*-gene cluster of a type IV secretion system, which is frequently found in *Alphaproteobacteria* megaplasmids [63], and 5 putative *dsb*-like genes the products of which may be involved in introducing disulfide bonds into diverse substrate proteins [64].

As mentioned above, the removal of pAMI1 from JCM 7686 cells influenced colony morphology and affected the growth mode in rich LB medium (but not in minimal-salt AC medium) (Additional file 13). Analysis of the pAMI1 sequence revealed that 40% of its genes encode proteins involved in amino acid transport and metabolism. Since amino acids serve as the main nitrogen source in LB medium, while nitrogen is provided by ammonium ions in AC medium, this may explain the observed growth differences.

Prophages and other prophage-related regions

In the *P. aminophilus* chromosome, we identified 10 regions encoding phage-related proteins (Figure 1). These constitute 8.2% of the host chromosome. Close inspection revealed that only 6 of them encode a full set of proteins crucial for the phage "life cycle". Thus only these regions were considered to represent putative prophages and designated ϕPam-1 to ϕPam-6, respectively (Figure 3A). The predicted sizes of the identified prophages range from 32.9 to 43.9 kb and they comprise between 41 and 49 phage-related genes coding for proteins involved in integration, replication, packaging, capsid and tail assembly, and lysis. For four of the prophages we could identify putative integration sites, which were tRNA sequences (tRNA-Pro for ϕPam-1, tRNA-Arg for ϕPam-3, tRNA-Gly for ϕPam-5 and tRNA-Met for ϕPam-6).

The vast majority of genes annotated within the prophage regions encode hypothetical proteins of unknown function. However, since prophages may encode fitness factors (e.g. virulence genes or metabolic modules) for their lysogenic hosts (e.g. [65]) we performed in-depth homology searches and found that genes JCM7686_0776 and JCM7686_1223 (within ϕPam-1 and ϕPam-2) encode tellurite resistance proteins (TerB). Moreover, we found that all but one (ϕPam-3) of the prophages encode methyltransferases (Figure 3A).

The JCM 7686 prophages exhibit high reciprocal nucleotide sequence similarity (Figure 3A), but they share only limited and localized homology with other prophages. This finding suggests their uniqueness among so far identified bacteriophages. Comparative analyses of the phage structural proteins revealed that five of the JCM 7686 prophages (ϕPam-2-ϕPam-6) encode major capsid proteins similar to those of HK97-like phages (*Siphoviridae*), which may suggest some evolutionary relationship.

To identify functional phages we applied the method of aggressive induction with mitomycin C. Using this procedure, only one of the JCM 7686 prophages, ϕPam-6, was induced. The ϕPam-6 prophage is flanked in the genome by 17-bp direct repeats (5′-CCCTCCTCCGCTACCAT-3′) which were recognized as the phage attachment site. In addition, ϕPam-6 encodes a serine family recombinase, rather than a tyrosine family integrase that is typical for phages (including the other JCM 7686 prophages). Following induction, bacteriophage particles were visualized by transmission electron microscopy using negative staining with uranyl acetate (Additional file 16) and it was confirmed by the restriction analysis that they contain ϕPam-6 DNA. Hence, ϕPam-6 is the first functional phage identified in *Paracoccus* spp.

Among the predicted prophage-related regions we also identified one gene transfer agent (GTA) cluster of 14.7 kb, which contains 18 putative genes (Figure 3B). This GTA cluster is located upstream of the *cysE* gene, encoding serine O-acetyltransferase (involved in cysteine biosynthesis), which is a common location for GTAs in bacterial genomes [66]. The identified region shares synteny and exhibits 80% nucleotide sequence identity with a GTA cluster found within chromosome II of *Paracoccus denitrificans* PD1222 [GenBank:CP000490]. Moreover, the *P. aminophilus* GTA exhibits homology to other RcGTA-like gene clusters (the archetype of this group is *Rhodobacter capsulatus* GTA) identified in various representatives of *Alphaproteobacteria* [66].

Transposable elements (TEs)

Analysis of the JCM 7686 genome identified 72 predicted transposase genes (Table 1); however, 32 of them are truncated (Additional file 17). The encoded transposases were classified and particular TEs distinguished (Additional file 17).

We identified 11 types of complete insertion sequences representing 5 IS families [IS*3* (IS*407*, IS*51* groups), IS*5* (IS*427*, IS*903* groups), IS*66*, IS*110* and IS*1182*] (Additional file 17). Most of the identified ISs are present as a single copy in the JCM 7686 genome. The exceptions are IS*Pam1* and *ISPam2* (both have 2 copies), as well as IS*Pam5* (8 copies). As shown in Figure 3C, the majority of the identified ISs contain a single ORF encoding a transposase (members of the IS*903* group of the IS*5* family, IS*110* and IS*1182* families) or carry two overlapping ORFs and possess a conserved frame-shift motif (members of the IS*407* group of the IS*3* family and IS*427* group of the IS*5* family). The frame-shift sequences are most probably responsible for the generation of a fusion protein (ORF1 + ORF2) as a result of programmed translational frameshifting [67]. A putative trans-frame transposase may also be potentially produced by IS*Pam8* of the IS*66* family. IS*Pam8* also contains two additional ORFs (Figure 3C), which may be involved in the regulation of transposition.

Within the plasmid component of the JCM 7686 genome we also identified a non-composite, cryptic transposon Tn*3434a* of the Tn*3* family. It is present in 4 copies: in pAMI1, pAMI7 and pAMI8 (2 copies). The presence of two identical divergently oriented copies of Tn*3434a* within pAMI8 resulted in inversion of an

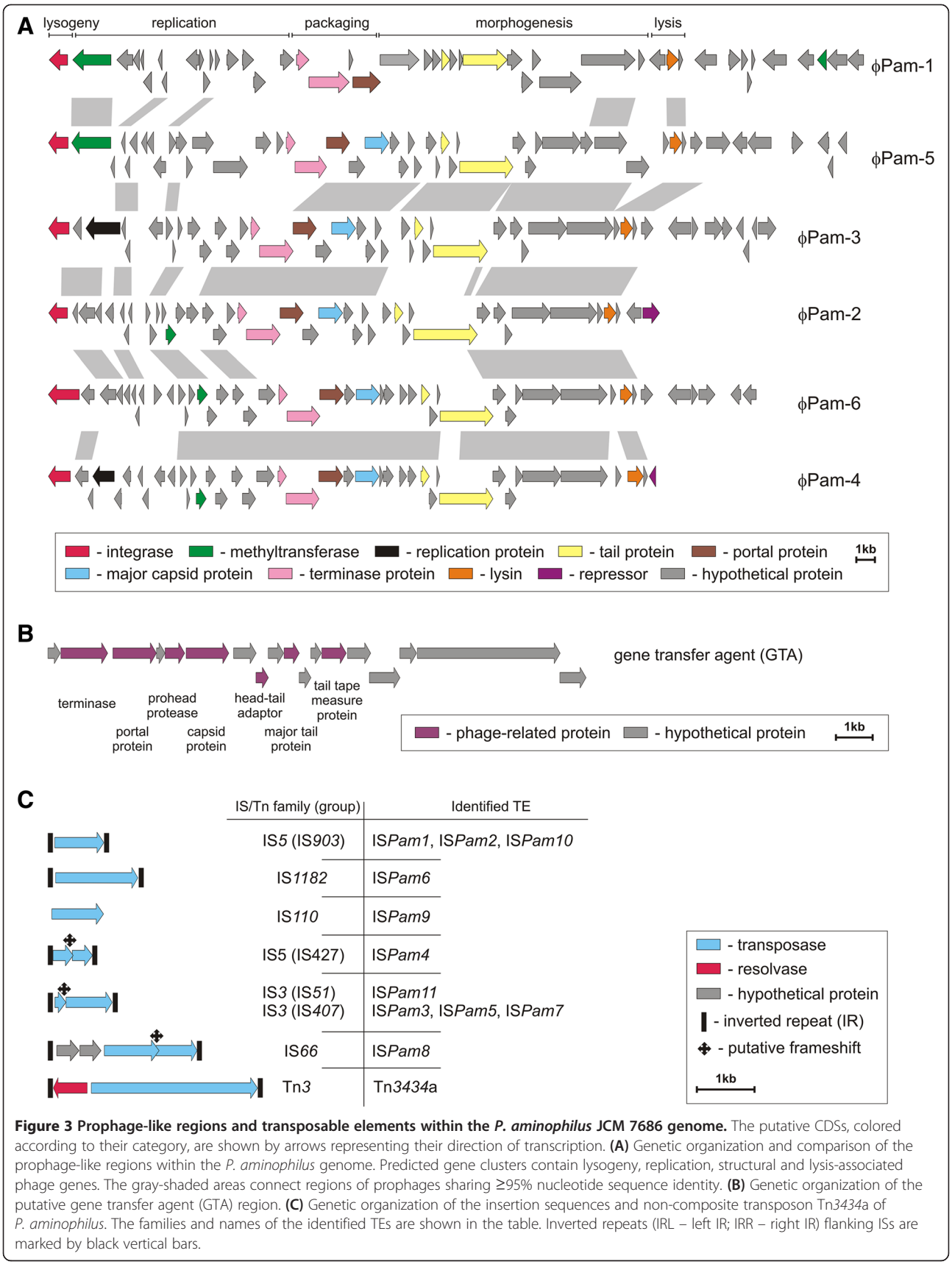


Figure 3 Prophage-like regions and transposable elements within the *P. aminophilus* JCM 7686 genome. The putative CDSs, colored according to their category, are shown by arrows representing their direction of transcription. **(A)** Genetic organization and comparison of the prophage-like regions within the *P. aminophilus* genome. Predicted gene clusters contain lysogeny, replication, structural and lysis-associated phage genes. The gray-shaded areas connect regions of prophages sharing ≥95% nucleotide sequence identity. **(B)** Genetic organization of the putative gene transfer agent (GTA) region. **(C)** Genetic organization of the insertion sequences and non-composite transposon Tn*3434*a of *P. aminophilus*. The families and names of the identified TEs are shown in the table. Inverted repeats (IRL – left IR; IRR – right IR) flanking ISs are marked by black vertical bars.

approx. 15-kb DNA region due to site-specific resolvase-mediated recombination.

In a previous study, using a trap plasmid strategy, we showed that IS*Pam1*, IS*Pam2*, IS*Pam3* and Tn*3434a* are fully functional elements [17]. However, such analyses have some limitations because, although it may identify most active elements, other functional TEs can be missed due to their low transposition frequency.

Comparative genomics of *Paracoccus* spp

Until now, the annotated genome sequences of only four *Paracoccus* spp. strains are available in public databases: *P. denitrificans* PD1222 ([GenBank:CP000489], [GenBank:CP000490] and [GenBank:CP000491]), *Paracoccus* sp. strain TRP [7], *P. denitrificans* SD1 [8] and *Paracoccus* sp. N5 ([GenBank:NZ_AQUO01000001], [GenBank:NZ_AQUO01000002] and [GenBank:NZ_AQUO01000003]). In this study we add the genome of *P. aminophilus* JCM 7686 and comparative analysis of five *Paracoccus* spp. genomes is now possible.

A total of 1001 genes (20 - 27% of all genes) can be considered as core *Paracoccus* genes, since they are present in all five analyzed genomes (Figure 4). Functional classification of the proteins encoded by these genes showed that, in a vast majority of cases, they belong to five COG groups (J, O, C, F, H) that represent basic housekeeping functions, expected to be encoded on chromosomes, and therefore conserved among bacterial genomes.

Comparative genomic analyses disclosed some specific features of the *P. aminophilus* genome. The chromosome of strain JCM 7686 contains 757 singletons (22% of all genes) that are mainly located within prophages, although other gene clusters are also included, e.g. a type III secretion system with 19 unique genes. The remaining singleton genes are encoded by the plasmids (46% of all plasmid genes). A relatively large portion of singletons within the genome is a feature of all the analyzed *Paracoccus* genomes, which suggests that these strains are not that closely related.

We also performed a complex phylogenetic analysis of *P. aminophilus* JCM 7686 in relation to other fully sequenced members of the class *Alphaproteobacteria*. A phylogenetic tree based on 453 core genes was computed (Figure 5). *P. aminophilus* JCM 7686 is clustered together with the other completely sequenced *Paracoccus* spp. strains, but it forms an outgroup, which indicates its "distant" relationship to representatives of other species (Figure 5).

Conclusions

Chromids have been found in bacteria belonging to many different phyla, including *Actinobacteria, Bacteroidetes, Chloroflexi, Cyanobacteria, Deinococcus-Thermus*, *Firmicutes*, *Proteobacteria* and *Spirochaetes* [60,61]. It was shown that chromids of bacteria classified to separate phylogenetic groups carry replication systems of different types, which indicates that these replicons were derived from plasmid precursors typical for particular groups of hosts, and were formed independently in the course of evolution [60,61].

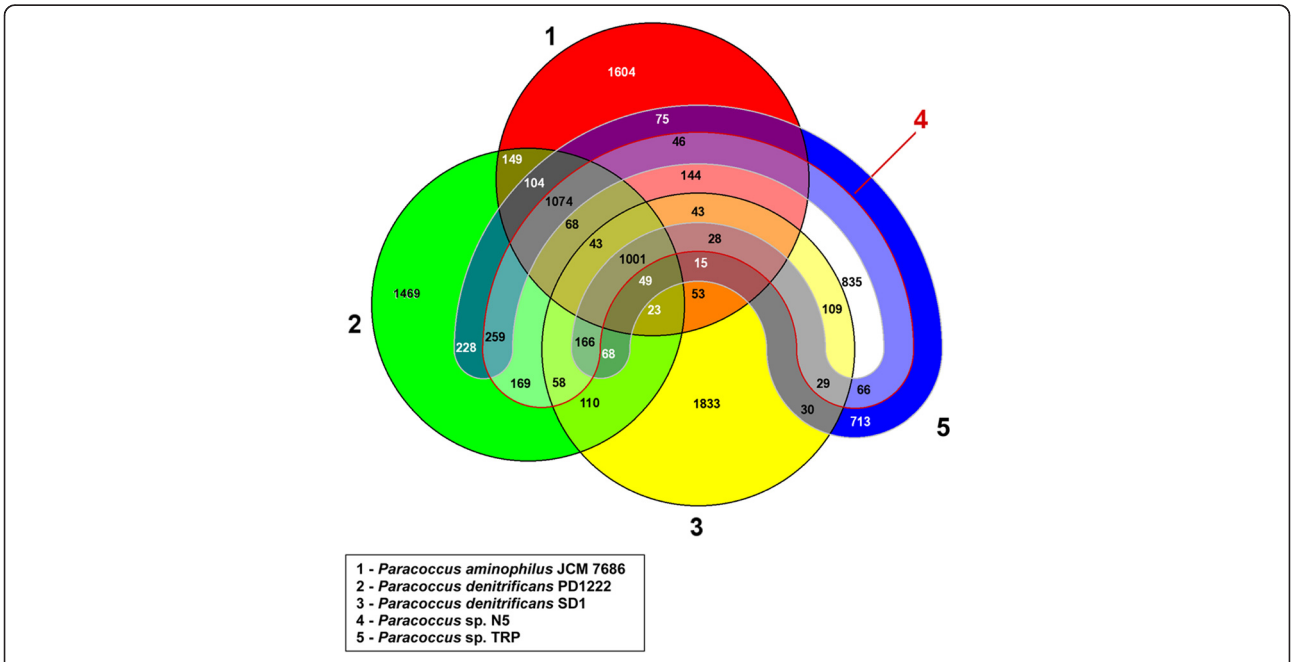


Figure 4 Venn diagram of the comparison of five sequenced and annotated *Paracoccus* genomes. The differently colored areas represent the compared strains. The central area describes the core genome consisting of 1001 genes.

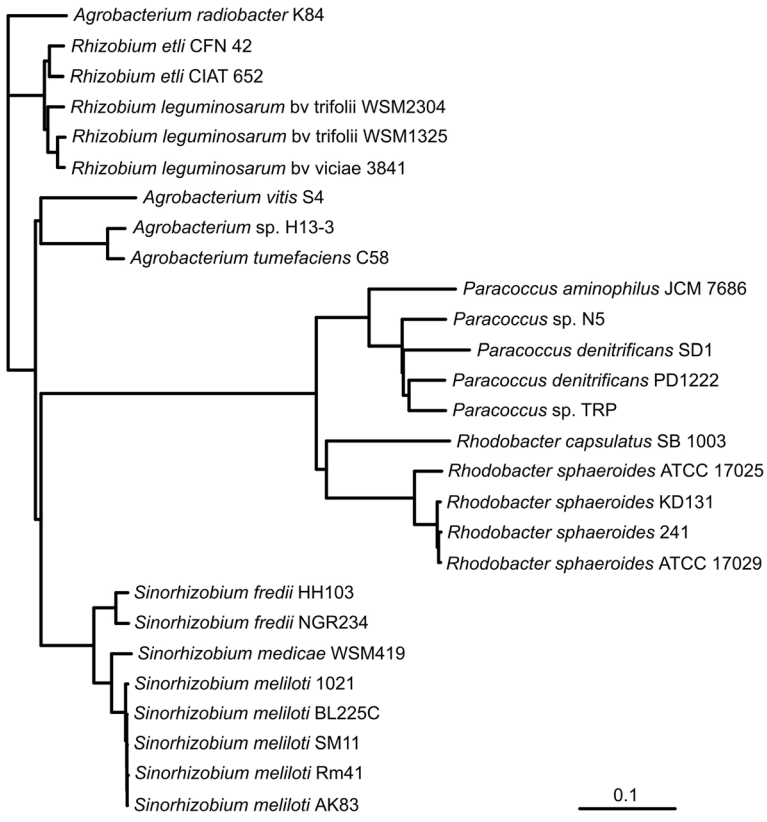


Figure 5 Phylogenetic tree based on 453 core genes of the selected strains of *Alphaproteobacteria*. Multiple sequence alignments of concatenated core gene sequences were calculated using EDGAR tool. Plasmid sequences were included.

Chromids share a number of major characteristics: (i) considerable size (they are usually the second largest replicons in the cell, being larger than the chromosomes of some bacteria), (ii) the presence of plasmid-type replication systems, (iii) G + C nucleotide contents comparable to that of chromosomes, (iv) codon usage similar to that of chromosomes (points iii and iv indicate long-term co-evolution of chromosomes and chromids), (v) the presence of housekeeping genes, typical for chromosomes (removal of chromids from the bacterial cell causes a lethal effect), and (vi) the presence of adaptive genes, typical for plasmids, enabling adaptation of bacteria to a given ecological niche [60,68,69]. A typical feature of chromids is their high variability, resulting mainly from the low density of housekeeping genes. Therefore, they are considered to be specific "training sites" in which different evolutionary variants are "tested" [68,70,71], which leads to structural variability of multireplicon genomes in different strains of the same genus [72].

The formation of multi-replicon genomes, in which the basic genomic information is split between various replicons, seems to be beneficial to bacteria. The presence of additional replicons enables more rapid duplication of the genetic information, and permits the maintenance of a larger genome while keeping a high rate of cell division. Moreover, it was shown that the frequency of dimer generation by bacterial chromosomes increases exponentially in relation to their size; therefore, the reduction of chromosome size (by dividing the genetic information essential for host viability, into two or more replicons) allows minimization of this phenomenon [73].

Many *Paracoccus* spp. genomes also possess a multi-replicon structure. A good example of such a genome architecture is *P. aminophilus* JCM 7686, which carries one chromosome and eight plasmid-like replicons. Our *in silico* analyses revealed that pAMI1, pAMI4, pAMI5, pAMI6 and pAMI8 carry many genes (also predicted core genes) showing homology to ORFs within *P. denitrificans* PD1222 chromosomes (Additional file 11), thus they may be considered as putative chromids. However, relying only on the bioinformatic studies, the concept of essentiality of particular replicons is highly speculative, thus to define their nature we applied additional analyses.

In this study, acknowledging the importance of functional analyses, we have shown that only pAMI5 and pAMI6 should be classified as chromids, since their presence in the host cells is essential for their proper functioning. pAMI5 could not be removed from the host cells at all, and while it was possible to obtain a pAMI6-

less derivative, this strain was unable to grow in minimal media, i.e. in conditions similar to those in the natural environment. Based on this observation we propose the classification of bacterial chromids into two types: "primary" chromids (e.g. pAMI5), which are indispensable for host viability and carry genes of the core genome, thus their elimination from the host cells is impossible under any environmental conditions, and "secondary" chromids (e.g. pAMI6) which were probably formed quite recently from an evolutionary point of view (e.g. they contain different REP systems, and have more characteristics typical for plasmids). Moreover, secondary chromids seems to carry a species (or genera) -specific pool of genes that are crucial for survival in the natural environment, but are not essential under "safe" laboratory conditions, thus the replicons are only "facultatively" essential [74].

According to this definition, the chromosome II of *P. denitrificans* PD1222 should be reclassified as a primary chromid. It is noteworthy that chromosome II of PD1222 and pAMI5 carry related *dnaA*-like replication systems, which may be a typical feature of the primary chromids of *Paracoccus* spp. This may facilitate further classifications of such replicons in other representative of this genus.

Methods

Strains, plasmids and culture conditions

The strains used in this study are presented in Additional file 18. They were grown in Luria-Bertani (LB) medium [75], TY medium (*Rhizobium etli* CE3) [76] and the minimal salts medium AC [77] at 37°C (*E. coli*) and 30°C (other strains). Where necessary, the media were supplemented with arabinose (0.2%), glucose (1%), kanamycin (50–1000 μg/ml) and rifampicin (50 μg/ml). The plasmids used and constructed in this study are described in Additional file 19.

DNA isolation, standard genetic manipulations and introduction of plasmid DNA into bacterial cells

The isolation of total DNA and plasmids, as well as common DNA manipulation methods were performed as described by Sambrook and Russell (2001) [75]. The visualization of mega-sized replicons was achieved by in-gel lysis and DNA electrophoresis according to a method described by Wheatcroft et al. (1990) [78]. PCR was performed in a Mastercycler using HiFi polymerase (Qiagen; with supplied buffer), dNTP mixture, total DNA (or plasmid DNA) of *P. aminophilus* as the template and appropriate oligonucleotide primers (Additional file 20). Transformation of *E. coli* strains was performed according to the method of Kushner [79]. Triparental mating was performed as previously described [80].

Determination of the sequence specificity of methyltransferases using the endonuclease protection assay

Putative MTase genes were PCR-amplified from *P. aminophilus* genomic DNA using specific primer pairs (Additional file 20). The amplified genes were cloned into expression vectors pET-28a or pET-30a (Additional file 19). Obtained plasmids were used to transform *E. coli* ER2566 and the strains were propagated for expression. To repress T7 RNA polymerase expression in the ER2566 strains before induction, glucose was added to cultures at a final concentration of 1%. The DNAs of recombinant plasmids isolated from IPTG-induced and non-induced bacterial cultures were used as substrates for cleavage by selected restriction enzymes.

Phage excision and transmission electron microscopy

Mitomycin C (0.5 mg/ml) was added to a logarithmic phase culture of *P. aminophilus* in 20 ml LB medium. Following incubation for a further 8 h, the culture was centrifuged at 11,000 g for 10 min to pellet the cells and the supernatant fraction was passed through a membrane filter (0.22 μm pore size). Phage particles were collected by centrifugation (2 h at 100,000 g) and resuspended in 50 μl of SM buffer (50 mM Tris–HCl, pH 7.5, 100 mM NaCl, 8 mM $MgSO_4 \times 7H_2O$, 0.01% gelatin). One drop of the suspension was applied to the surface of a Formvar-coated grid, negatively stained with 2% uranyl acetate [81], and then examined using a LEO 912AB transmission electron microscope.

DNA sequencing

Total genomic DNA was isolated from an overnight culture of *P. aminophilus* JCM 7686 using the DNeasy Blood & Tissue Kit (Qiagen). This was used for the construction of NGS (Next Generation Sequencing) libraries: (i) GS FLX + shotgun library [using the GS FLX + library preparation kit (Roche)], (ii) an 8-kb-long paired-end library [using the GS FLX Paired-End kit (Roche)] and (iii) Illumina Paired-End library [using the Illumina TruSeq2.0 kit (Illumina Inc.)]. The genomic libraries were sequenced using a Genome Sequencer FLX + System (Roche) and Illumina HiScanSQ Genome Analyzer (Illumina Inc.) in the DNA Sequencing and Oligonucleotide Synthesis Laboratory (oligo.pl) at the Institute of Biochemistry and Biophysics, Polish Academy of Sciences. The sequence data was assembled into contigs and scaffolds using Newbler De Novo Assembler (454 Sequencing System Software, Roche). Any remaining gaps were closed using the Expand Long Template PCR System (Roche), followed by Sanger sequencing with an ABI3500 Genetic Analyzer (Life Technologies) using BigDye Terminator Mix v. 3.1 chemistry (Life Technologies).

Bioinformatics

Automatic gene prediction and annotation of the JCM 7686 genome were performed using GenDB 2.0 [82]. Automatic annotations and intergenic regions were analyzed and corrected manually by means of BLAST programs [83] and the PRIAM tool [84]. Putative tRNA genes were identified with the tRNAscan-SE program [85]. Each gene was functionally classified by assigning a Cluster of Orthologous Groups (COG) number and its corresponding COG category [86]. Comparison searches for insertion sequences were performed with ISfinder [87]. Finally, the Artemis software was used to visualize the genome [88].

Comparisons between the JCM 7686 genome and the genomes of other representatives of the *Alphaproteobacteria* (i.e. synteny analyses, identification of orthologous genes and classification of genes as core genes or singletons) were performed using the EDGAR tool [89]. It was also used for the creation of the phylogenetic tree. To construct the tree, 453 core genes from 27 genomes of the *Alphaproteobacteria* were computed. The multiple alignments for all core genes were created using MUSCLE [90]. Non matching parts of the alignments were masked using GBLOCKS [91] and removed subsequently. The remaining parts of all alignments were concatenated to one multiple alignment, which then was used to generate the phylogenetic tree applying PHYLIP [92]. Other comparative analyses were performed with the GeneOrder4.0 tool [93].

Nucleotide sequence accession numbers

The nucleotide sequences of *P. aminophilus* JCM 7686 chromosome and extrachromosomal replicons pAMI1-8 have been annotated and deposited in GenBank (NCBI) with respective accession numbers: CP006650, CP006651, GQ410978, GQ468939, CP006652, CP006653, CP006654, GQ468938, CP006655.

Additional files

Additional file 1: COG categories of the proteins of *P. aminophilus* JCM 7686.

Additional file 2: Diversity and distribution of 217 transcriptional regulators of *P. aminophilus* JCM 7686.

Additional file 3: Two-component systems and histidine kinases encoded by the *P. aminophilus* JCM 7686 genome.

Additional file 4: Toxin-antitoxin systems encoded by the *P. aminophilus* JCM 7686 genome.

Additional file 5: Chaperones and co-chaperonins encoded by the *P. aminophilus* JCM 7686 genome.

Additional file 6: Genes encoding DNA repair related proteins within the *P. aminophilus* JCM 7686 genome.

Additional file 7: Restriction patterns of *P. aminophilus* JCM 7686 genomic DNA cleaved with selected restriction endonucleases showing the protection of GANTC sites by the CcrM methylase (JCM7685_3079). ND - undigested DNA. M – GeneRuler 100–10,000 bp size marker.

Additional file 8: Summary of the sensitivity of various restriction endonucleases to DNA modifications introduced by the JCM7686_0772 and JCM7686_2655 proteins (m^5C MTases).

Additional file 9: Summary of the sensitivity of various restriction endonucleases to DNA modifications introduced by the JCM7686_0815 protein (m^4C MTase).

Additional file 10: Summary of the sensitivity of various restriction endonucleases to DNA modifications introduced by the JCM7686_1231, JCM7686_2255 and JCM7686_2934 proteins (m^6A MTases).

Additional file 11: The number of homologous proteins encoded by JCM 7686 plasmids and *P. denitrificans* PD1222 chromosomes. The analysis was performed using the GeneOrder 4.0 tool. Proteins were considered homologous only if the BLAT threshold scores were >200. The results were verified manually by BLAST comparisons.

Additional file 12: Plasmid profiles of the JCM 7686 wild-type strain and its derivatives deprived of particular megasized-replicons.

Additional file 13: The growth mode of wild-type JCM 7686 and the pAMI1-less derivative in liquid LB medium.

Additional file 14: Host ranges of *P. aminophilus* JCM 7686 pAMI1-8 replicons.

Additional file 15: Schematic diagram of *de novo* NAD biosynthesis from aspartate to nicotinic acid mononucleotide (NaMN).

Additional file 16: Transmission electron micrograph of tailed bacteriophage ϕPam-6 of *P. aminophilus* JCM 7686.

Additional file 17: Transposase genes within the *P. aminophilus* JCM 7686 genome.

Additional file 18: Bacterial strains used in this study.

Additional file 19: Plasmids used and constructed in this study.

Additional file 20: Oligonucleotide primers used in this study.

Abbreviations

CDS: Coding sequence; COG: Cluster of Orthologous Group; DMF: *N,N*-dimethylformamide; GTA: Gene transfer agent; HPK: Histidine protein kinase; IR: Inverted repeat; IS: Insertion sequence; LB: Luria-bertani; MGE: Mobile genetic elements; MTase: methyltransferase; NAD: Nicotinamide adenine dinucleotide; NaMN: Nicotinic acid mononucleotide; NCBI: National Center for Biotechnology Information; ORF: Open reading frame; PCR: Polymerase chain reaction; TE: Transposable element; TMo: Transposable module; Tn: Transposon.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

LD and DB conceived and designed the experiments; LD coordinated the project; LD, JC, DW, MR, PM and MS performed the experiments; LD, JC, MR and DW analyzed the data; LD, DB, AS and AP contributed reagents/materials/analysis tools; DB, AS and AP supervised the work; LD, DB and DW wrote the paper. All authors read and approved the final manuscript.

Acknowledgements

We would like to thank J. Baj for scientific advice, as well as L. Kowalski, K. Kuzmicz, K. Romaniuk and A. Szych for technical assistance. We are also grateful to J. Gawor from the DNA Sequencing and Oligonucleotide Synthesis Laboratory (oligo.pl; Institute of Biochemistry and Biophysics, Polish Academy of Sciences) for his excellent work during the genome sequencing. The work was supported by the Ministry of Science and Higher Education, Poland (grants: IP2010 008670 and IP2011 011471).

Author details

[1]Department of Bacterial Genetics, Institute of Microbiology, Faculty of Biology, University of Warsaw, Miecznikowa 1, 02-096 Warsaw, Poland. [2]Institute for Genome Research and Systems Biology, Center for Biotechnology, Bielefeld University, D-33594 Bielefeld, Germany. [3]Department

of Virology, Institute of Microbiology, Faculty of Biology, University of Warsaw, Miecznikowa 1, 02-096 Warsaw, Poland.

Received: 9 September 2013 Accepted: 7 February 2014
Published: 12 February 2014

References

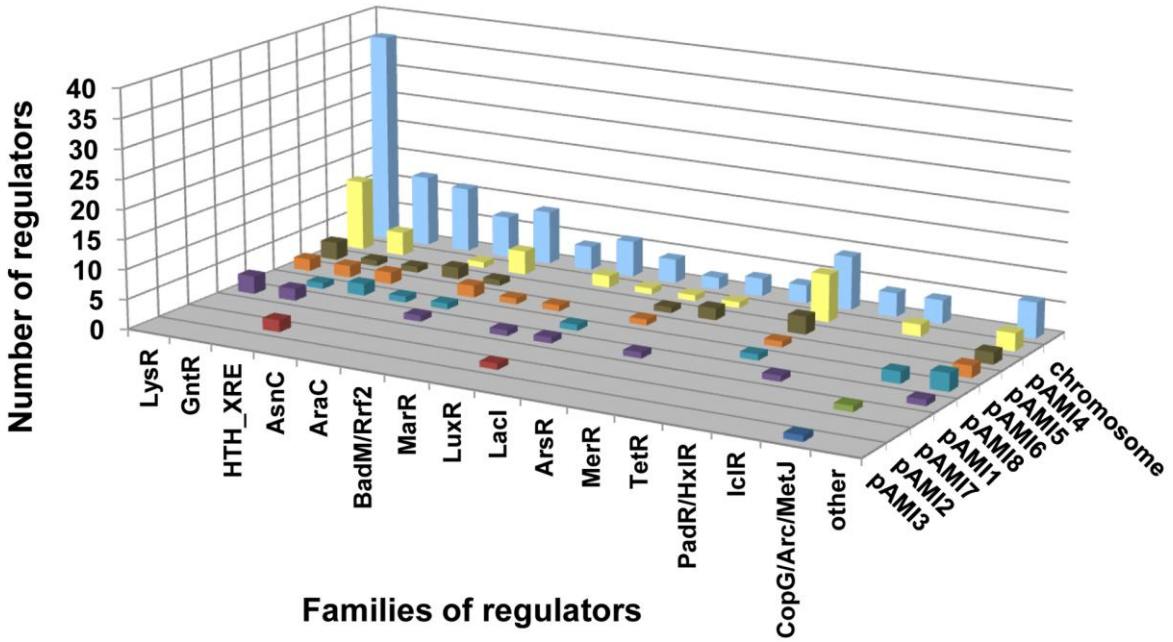
1. Baker SC, Ferguson SJ, Ludwig B, Page MD, Richter OM, van Spanning RJ: **Molecular genetics of the genus *Paracoccus*: metabolically versatile bacteria with bioenergetic flexibility.** *Microbiol Mol Biol Rev* 1998, **62**(4):1046–1078.
2. Katayama Y, Hiraishi A, Kuraishi H: ***Paracoccus thiocyanatus* sp. nov., a new species of thiocyanate-utilizing facultative chemolithotroph, and transfer of *Thiobacillus versutus* to the genus *Paracoccus* as *Paracoccus versutus* comb. nov. with emendation of the genus.** *Microbiology* 1995, **141:**1469–1477.
3. Kelly DP, Euzeby JP, Goodhew CF, Wood AP: **Redefining *Paracoccus denitrificans* and *Paracoccus pantotrophus* and the case for a reassessment of the strains held by international culture collections.** *Int J Syst Evol Microbiol* 2006, **56:**2495–2500.
4. Kumaraswamy R, Sjollema K, Kuenen G, van Loosdrecht M, Muyzer G: **Nitrate-dependent [Fe(II)EDTA]2- oxidation by *Paracoccus ferrooxidans* sp. nov., isolated from a denitrifying bioreactor.** *Syst Appl Microbiol* 2006, **29:**276–286.
5. Doronina NV, Trotsenko YA, Krausova VI, Suzina NE: ***Paracoccus methylutens* sp. nov. - a new aerobic facultatively methylotrophic bacterium utilizing dichloromethane.** *Syst Appl Microbiol* 1998, **21:**230–236.
6. Sun M, Luo Y, Teng Y, Christie P, Jia Z, Li Z: **Tenax TA extraction to understand the rate-limiting factors in methyl-beta-cyclodextrin-enhanced bioremediation of PAH-contaminated soil.** *Biodegradation* 2013, **24:**365–375.
7. Li K, Wang S, Shi Y, Qu J, Zhai Y, Xu L, Xu Y, Song J, Liu L, Rahman MA, *et al*: **Genome sequence of *Paracoccus* sp. Strain TRP, a chlorpyrifos biodegrader.** *J Bacteriol* 2011, **193**(7):1786–1787.
8. Siddavattam D, Karegoudar TB, Mudde SK, Kumar N, Baddam R, Avasthi TS, Ahmed N: **Genome of a novel isolate of *Paracoccus denitrificans* capable of degrading *N,N*-dimethylformamide.** *J Bacteriol* 2011, **193**(19):5598–5599.
9. Baj J, Piechucka E, Bartosik D, Wlodarczyk M: **Plasmid occurrence and diversity in the genus *Paracoccus*.** *Acta Microbiol Pol* 2000, **49**(3–4):265–270.
10. Bartosik D, Putyrski M, Dziewit L, Malewska E, Szymanik M, Jagiello E, Lukasik J, Baj J: **Transposable modules generated by a single copy of insertion sequence IS*Pme1* and their influence on structure and evolution of natural plasmids of *Paracoccus methylutens* DM12.** *J Bacteriol* 2008, **190**(9):3306–3313.
11. Bartosik D, Baj J, Sochacka M, Piechucka E, Wlodarczyk M: **Molecular characterization of functional modules of plasmid pWKS1 of *Paracoccus pantotrophus* DSM 11072.** *Microbiology* 2002, **148**(Pt 9):2847–2856.
12. Bartosik D, Baj J, Wlodarczyk M: **Molecular and functional analysis of pTAV320, a *repABC*-type replicon of the *Paracoccus versutus* composite plasmid pTAV1.** *Microbiology* 1998, **144:**3149–3157.
13. Bartosik D, Witkowska M, Baj J, Wlodarczyk M: **Characterization and sequence analysis of the replicator region of the novel plasmid pALC1 from *Paracoccus alcaliphilus*.** *Plasmid* 2001, **45**(3):222–226.
14. Bartosik D, Baj J, Bartosik AA, Wlodarczyk M: **Characterization of the replicator region of megaplasmid pTAV3 of *Paracoccus versutus* and search for plasmid-encoded traits.** *Microbiology* 2002, **148**(Pt 3):871–881.
15. Bartosik D, Sochacka M, Baj J: **Identification and characterization of transposable elements of *Paracoccus pantotrophus*.** *J Bacteriol* 2003, **185**(13):3753–3763.
16. Bartosik D, Szymanik M, Baj J: **Identification and distribution of insertion sequences of *Paracoccus solventivorans*.** *Appl Environ Microbiol* 2003, **69**(12):7002–7008.
17. Dziewit L, Baj J, Szuplewska M, Maj A, Tabin M, Czyzkowska A, Skrzypczyk G, Adamczuk M, Sitarek T, Stawinski P, *et al*: **Insights into the transposable mobilome of *Paracoccus* spp. (*Alphaproteobacteria*).** *PLoS One* 2012, **7**(2):e32277.
18. Mikosa M, Sochacka-Pietal M, Baj J, Bartosik D: **Identification of a transposable genomic island of *Paracoccus pantotrophus* DSM 11072 by its transposition to a novel entrapment vector pMMB2.** *Microbiology* 2006, **152**(Pt 4):1063–1073.
19. Szuplewska M, Bartosik D: **Identification of a mosaic transposable element of *Paracoccus marcusii* composed of insertion sequence IS*Pmar4* (IS*As1* family) and an IS*1247a*-driven transposable module (TMo).** *FEMS Microbiol Lett* 2009, **292**(2):216–221.
20. van Spanning RJ, de Boer AP, Slotboom DJ, Reijnders WN, Stouthamer AH: **Isolation and characterization of a novel insertion sequence element, IS*1248*, in *Paracoccus denitrificans*.** *Plasmid* 1995, **34**(1):11–21.
21. Urakami T, Araki H, Oyanagi H, Suzuki K, Komagata K: ***Paracoccus aminophilus* sp. nov. and *Paracoccus aminovorans* sp. nov., which utilize *N, N*-dimethylformamide.** *Int J Syst Bacteriol* 1990, **40**(3):287–291.
22. Dziewit L, Jazurek M, Drewniak L, Baj J, Bartosik D: **The SXT conjugative element and linear prophage N15 encode toxin-antitoxin-stabilizing systems homologous to the *tad-ata* module of the *Paracoccus aminophilus* plasmid pAMI2.** *J Bacteriol* 2007, **189**(5):1983–1997.
23. Dziewit L, Adamczuk M, Szuplewska M, Bartosik D: **DIY series of genetic cassettes useful in construction of versatile vectors specific for *Alphaproteobacteria*.** *J Microbiol Methods* 2011, **86**(2):166–174.
24. Dziewit L, Kuczkowska K, Adamczuk M, Radlinska M, Bartosik D: **Functional characterization of the type II PamI restriction-modification system derived from plasmid pAMI7 of *Paracoccus aminophilus* JCM 7686.** *FEMS Microbiol Lett* 2011, **324**(1):56–63.
25. Dziewit L, Dmowski M, Baj J, Bartosik D: **Plasmid pAMI2 of *Paracoccus aminophilus* JCM 7686 carries *N, N*-dimethylformamide degradation-related genes whose expression is activated by a LuxR family regulator.** *Appl Environ Microbiol* 2010, **76**(6):1861–1869.
26. Studholme DJ, Buck M: **The biology of enhancer-dependent transcriptional regulation in bacteria: insights from genome sequences.** *FEMS Microbiol Lett* 2000, **186:**1–9.
27. Martinez-Antonio A, Collado-Vides J: **Identifying global regulators in transcriptional regulatory networks in bacteria.** *Curr Opin Microbiol* 2003, **6:**482–489.
28. Maddocks SE, Oyston PC: **Structure and function of the LysR-type transcriptional regulator (LTTR) family proteins.** *Microbiology* 2008, **154:**3609–3623.
29. Subramoni S, Venturi V: **LuxR-family 'solos': bachelor sensors/regulators of signalling molecules.** *Microbiology* 2009, **155**(Pt 5):1377–1385.
30. Wolanin PM, Thomason PA, Stock JB: **Histidine protein kinases: key signal transducers outside the animal kingdom.** *Genome Biol* 2002, **3**(10). REVIEWS3013.
31. Drepper T, Wiethaus J, Giaourakis D, Gross S, Schubert B, Vogt M, Wiencek Y, McEwan AG, Masepohl B: **Cross-talk towards the response regulator NtrC controlling nitrogen metabolism in *Rhodobacter capsulatus*.** *FEMS Microbiol Lett* 2006, **258:**250–256.
32. Wang X, Vu A, Lee K, Dahlquist FW: **CheA-receptor interaction sites in bacterial chemotaxis.** *J Mol Biol* 2012, **422:**282–290.
33. Lamarche MG, Wanner BL, Crepin S, Harel J: **The phosphate regulon and bacterial virulence: a regulatory network connecting phosphate homeostasis and pathogenesis.** *FEMS Microbiol Rev* 2008, **32:**461–473.
34. Ausmees N, Jacobs-Wagner C: **Spatial and temporal control of differentiation and cell cycle progression in *Caulobacter crescentus*.** *Annu Rev Microbiol* 2003, **57:**225–247.
35. Janausch IG, Zientz E, Tran QH, Kroger A, Unden G: **C4-dicarboxylate carriers and sensors in bacteria.** *Biochim Biophys Acta* 2002, **1553:**39–56.
36. Hughes DT, Clarke MB, Yamamoto K, Rasko DA, Sperandio V: **The QseC adrenergic signaling cascade in Enterohemorrhagic *E. coli* (EHEC).** *PLoS Pathog* 2009, **5:**e1000553.
37. Bird TH, Du S, Bauer CE: **Autophosphorylation, phosphotransfer, and DNA-binding properties of the RegB/RegA two-component regulatory system in *Rhodobacter capsulatus*.** *J Biol Chem* 1999, **274:**16343–16348.
38. Jung K, Veen M, Altendorf K: **K + and ionic strength directly influence the autophosphorylation activity of the putative turgor sensor KdpD of *Escherichia coli*.** *J Biol Chem* 2000, **275:**40142–40147.
39. Jourlin C, Bengrine A, Chippaux M, Mejean V: **An unorthodox sensor protein (TorS) mediates the induction of the tor structural genes in response to trimethylamine N-oxide in *Escherichia coli*.** *Mol Microbiol* 1996, **20:**1297–1306.
40. Boor KJ: **Bacterial stress responses: what doesn't kill them can make then stronger.** *PLoS Biol* 2006, **4:**e23.
41. Butala M, Zgur-Bertok D, Busby SJ: **The bacterial LexA transcriptional repressor.** *Cell Mol Life Sci* 2009, **66:**82–93.
42. Ferullo DJ, Lovett ST: **The stringent response and cell cycle arrest in *Escherichia coli*.** *PLoS Genet* 2008, **4:**e1000300.

43. Christensen SK, Mikkelsen M, Pedersen K, Gerdes K: **RelE, a global inhibitor of translation, is activated during nutritional stress.** *Proc Natl Acad Sci U S A* 2001, **98:**14328–14333.
44. Lund PA: **Microbial molecular chaperones.** *Adv Microb Physiol* 2001, **44:**93–140.
45. Wagner EG: **Cycling of RNAs on Hfq.** *RNA Biol* 2013, **10:**619–626.
46. Chen B, Zhong D, Monteiro A: **Comparative genomics and evolution of the HSP90 family of genes across all kingdoms of organisms.** *BMC Genomics* 2006, **7:**156.
47. Martins-Pinheiro M, Marques RC, Menck CF: **Genome analysis of DNA repair genes in the alpha proteobacterium *Caulobacter crescentus*.** *BMC Microbiol* 2007, **7:**17.
48. Wigley DB: **Bacterial DNA repair: recent insights into the mechanism of RecBCD, AddAB and AdnAB.** *Nat Rev Microbiol* 2013, **11:**9–13.
49. Menck CF: **Shining a light on photolyases.** *Nat Genet* 2002, **32:**338–339.
50. Rebeil R, Nicholson WL: **The subunit structure and catalytic mechanism of the *Bacillus subtilis* DNA repair enzyme spore photoproduct lyase.** *Proc Natl Acad Sci U S A* 2001, **98:**9038–9043.
51. Falnes PO, Johansen RF, Seeberg E: **AlkB-mediated oxidative demethylation reverses DNA damage in *Escherichia coli*.** *Nature* 2002, **419:**178–182.
52. Marinus MG, Casadesus J: **Roles of DNA adenine methylation in host-pathogen interactions: mismatch repair, transcriptional regulation, and more.** *FEMS Microbiol Rev* 2009, **33:**488–503.
53. Collier J: **Epigenetic regulation of the bacterial cell cycle.** *Curr Opin Microbiol* 2009, **12:**722–729.
54. Drozdz M, Piekarowicz A, Bujnicki JM, Radlinska M: **Novel non-specific DNA adenine methyltransferases.** *Nucleic Acids Res* 2012, **40**(5):2119–2130.
55. Bartosik D, Bialkowska A, Baj J, Wlodarczyk M: **Construction of mobilizable cloning vectors derived from pBGS18 and their application for analysis of replicator region of a pTAV202 mini-derivative of *Paracoccus versutus* pTAV1 plasmid.** *Acta Microbiol Pol* 1997, **46**(4):387–392.
56. Falnes PO, Rognes T: **DNA repair by bacterial AlkB proteins.** *Res Microbiol* 2003, **154:**531–538.
57. Lilic M, Jovanovic M, Jovanovic G, Savic DJ: **Identification of the CysB-regulated gene, *hslJ*, related to the *Escherichia coli* novobiocin resistance phenotype.** *FEMS Microbiol Lett* 2003, **224:**239–246.
58. Tremblay PL, Hallenbeck PC: **Ammonia-induced formation of an AmtB-GlnK complex is not sufficient for nitrogenase regulation in the photosynthetic bacterium *Rhodobacter capsulatus*.** *J Bacteriol* 2008, **190:**1588–1594.
59. Acosta-Cruz E, Wisniewski-Dye F, Rouy Z, Barbe V, Valdes M, Mavingui P: **Insights into the 1.59-Mbp largest plasmid of *Azospirillum brasilense* CBG497.** *Arch Microbiol* 2012, **194:**725–736.
60. Harrison PW, Lower RP, Kim NK, Young JP: **Introducing the bacterial 'chromid': not a chromosome, not a plasmid.** *Trends Microbiol* 2010, **18:**141–148.
61. Villasenor T, Brom S, Davalos A, Lozano L, Romero D, Los Santos AG: **Housekeeping genes essential for pantothenate biosynthesis are plasmid-encoded in *Rhizobium etli* and *Rhizobium leguminosarum*.** *BMC Microbiol* 2011, **11:**66.
62. Landeta C, Davalos A, Cevallos MA, Geiger O, Brom S, Romero D: **Plasmids with a chromosome-like role in rhizobia.** *J Bacteriol* 2011, **193:**1317–1326.
63. Christie PJ: **Type IV secretion: the *Agrobacterium* VirB/D4 and related conjugation systems.** *Biochim Biophys Acta* 2004, **1694:**219–234.
64. Depuydt M, Messens J, Collet JF: **How proteins form disulfide bonds.** *Antioxid Redox Signal* 2011, **15:**49–66.
65. Tinsley CR, Bille E, Nassif X: **Bacteriophages and pathogenicity: more than just providing a toxin?** *Microbes Infect* 2006, **8:**1365–1371.
66. Lang AS, Beatty JT: **Importance of widespread gene transfer agent genes in alpha-proteobacteria.** *Trends Microbiol* 2007, **15:**54–62.
67. Sekine Y, Eisaki N, Ohtsubo E: **Translational control in production of transposase and in transposition of insertion sequence IS*3*.** *J Mol Biol* 1994, **235**(5):1406–1420.
68. Galardini M, Pini F, Bazzicalupo M, Biondi EG, Mengoni A: **Replicon-dependent bacterial genome evolution: the case of *Sinorhizobium meliloti*.** *Genome Biol Evol* 2013, **5:**542–558.
69. Mazur A, Stasiak G, Wielbo J, Koper P, Kubik-Komar A, Skorupska A: **Phenotype profiling of *Rhizobium leguminosarum* bv. trifolii clover nodule isolates reveal their both versatile and specialized metabolic capabilities.** *Arch Microbiol* 2013, **195:**255–267.
70. Cooper VS, Vohr SH, Wrocklage SC, Hatcher PJ: **Why genes evolve faster on secondary chromosomes in bacteria.** *PLoS Comput Biol* 2010, **6:**e1000732.
71. Morrow JD, Cooper VS: **Evolutionary effects of translocations in bacterial genomes.** *Genome Biol Evol* 2012, **4:**1256–1262.
72. Mazur A, Stasiak G, Wielbo J, Kubik-Komar A, Marek-Kozaczuk M, Skorupska A: **Intragenomic diversity of *Rhizobium leguminosarum* bv. trifolii clover nodule isolates.** *BMC Microbiol* 2011, **11:**123.
73. Val ME, Skovgaard O, Ducos-Galand M, Bland MJ, Mazel D: **Genome engineering in *Vibrio cholerae*: a feasible approach to address biological issues.** *PLoS Genet* 2012, **8**(1):e1002472.
74. Petersen J, Frank O, Goker M, Pradella S: **Extrachromosomal, extraordinary and essential–the plasmids of the *Roseobacter* clade.** *Appl Microbiol Biotechnol* 2013, **97:**2805–2815.
75. Sambrook J, Russell DW: *Molecular Cloning: a laboratory manual*. New York: Cold Spring Harbor Laboratory Press; 2001.
76. Noel KD, Sanchez A, Fernandez L, Leemans J, Cevallos MA: ***Rhizobium phaseoli* symbiotic mutants with transposon Tn*5* insertions.** *J Bacteriol* 1984, **158**(1):148–155.
77. Wood AP, Kelly DP: **Heterotrophic growth of *Thiobacillus* A2 on sugars and organic acids.** *Arch Microbiol* 1977, **113**(3):257–264.
78. Wheatcroft R, McRae GD, Miller RW: **Changes in the *Rhizobium meliloti* genome and the ability to detect supercoiled plasmids during bacteroid development.** *Mol Plant Microbe Interact* 1990, **3:**9–17.
79. Kushner SR: **An improved method for transformation of *E. coli* with ColE1 derived plasmids**. In *Genetic Engineering*. Edited by Boyer HB, Nicosia S. Amsterdam: Elsevier/North-Holland; 1978:17–23.
80. Bartosik D, Szymanik M, Wysocka E: **Identification of the partitioning site within the *repABC*-type replicon of the composite *Paracoccus versutus* plasmid pTAV1.** *J Bacteriol* 2001, **183**(21):6234–6243.
81. Ackermann HW: **Basic phage electron microscopy.** *Methods Mol Biol* 2009, **501:**113–126.
82. Meyer F, Goesmann A, McHardy AC, Bartels D, Bekel T, Clausen J, Kalinowski J, Linke B, Rupp O, Giegerich R, *et al*: **GenDB - an open source genome annotation system for prokaryote genomes.** *Nucleic Acids Res* 2003, **31:**2187–2195.
83. Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ: **Gapped BLAST and PSI-BLAST: a new generation of protein database search programs.** *Nucleic Acids Res* 1997, **25**(17):3389–3402.
84. Claudel-Renard C, Chevalet C, Faraut T, Kahn D: **Enzyme-specific profiles for genome annotation: PRIAM.** *Nucleic Acids Res* 2003, **31:**6633–6639.
85. Lowe TM, Eddy SR: **tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence.** *Nucleic Acids Res* 1997, **25:**955–964.
86. Tatusov RL, Fedorova ND, Jackson JD, Jacobs AR, Kiryutin B, Koonin EV, Krylov DM, Mazumder R, Mekhedov SL, Nikolskaya AN, *et al*: **The COG database: an updated version includes eukaryotes.** *BMC Bioinformatics* 2003, **4:**41.
87. Siguier P, Perochon J, Lestrade L, Mahillon J, Chandler M: **ISfinder: the reference centre for bacterial insertion sequences.** *Nucleic Acids Res* 2006, **34:**D32–D36.
88. Carver T, Berriman M, Tivey A, Patel C, Bohme U, Barrell BG, Parkhill J, Rajandream MA: **Artemis and ACT: viewing, annotating and comparing sequences stored in a relational database.** *Bioinformatics* 2008, **24:**2672–2676.
89. Blom J, Albaum SP, Doppmeier D, Puhler A, Vorholter FJ, Zakrzewski M, Goesmann A: **EDGAR: a software framework for the comparative analysis of prokaryotic genomes.** *BMC Bioinformatics* 2009, **10:**154.
90. Edgar RC: **MUSCLE: multiple sequence alignment with high accuracy and high throughput.** *Nucleic Acids Res* 2004, **32**(5):1792–1797.
91. Talavera G, Castresana J: **Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments.** *Syst Biol* 2007, **56**(4):564–577.
92. Felsenstein J: **PHYLIP - phylogeny inference package (version 3.2).** *Cladistics* 1989, **5:**164–166.
93. Mahadevan P, Seto D: **Rapid pair-wise synteny analysis of large bacterial genomes using web-based GeneOrder4.0.** *BMC Res Notes* 2010, **3:**41.

doi:10.1186/1471-2164-15-124
Cite this article as: Dziewit *et al.*: Architecture and functions of a multipartite genome of the methylotrophic bacterium *Paracoccus aminophilus* JCM 7686, containing primary and secondary chromids. *BMC Genomics* 2014 **15**:124.

TABLE S1. COG categories of the proteins of *P. aminophilus* JCM 7686.

Proteins predicted functions (COG categories)	Chromosome	pAMI1	pAMI2	pAMI3	pAMI4	pAMI5	pAMI6	pAMI7	pAMI8
L: DNA replication, recombination and repair	126	4	7	2	17	5	2	5	37
K: transcription	202	15	2	1	43	29	17	1	15
J: translation, ribosomal structure and biogenesis	163	2	0	0	6	2	1	0	0
D: cell division and chromosome partitioning	21	1	1	0	2	1	1	1	3
O: posttranslational modification, protein turnover, chaperones	114	0	0	0	2	6	2	0	9
M: cell envelope biogenesis, outer membrane	139	2	0	0	10	5	2	0	3
N: cell motility and secretion	24	0	0	0	0	4	0	0	1
V: defense mechanisms	38	0	0	0	9	4	3	0	5
P: inorganic ion transport and metabolism	210	14	0	0	43	38	20	0	5
T: signal transduction mechanisms	89	4	0	0	6	14	7	0	8
U: intracellular trafficking and secretion	47	0	1	0	1	0	3	1	9
C: energy production and conversion	183	5	0	0	23	10	9	0	20
G: carbohydrate transport and metabolism	139	3	0	0	31	18	9	0	6
E: amino acid transport and metabolism	385	41	0	0	65	52	31	0	1
F: nucleotide transport and metabolism	75	2	0	0	4	3	5	0	0
H: coenzyme metabolism	124	3	0	0	8	9	8	0	2
I: lipid metabolism	105	5	0	0	19	14	10	0	2
Q: secondary metabolite biosynthesis, transport and catabolism	82	10	0	0	17	8	4	0	0
R: general function prediction only	316	16	0	1	51	28	18	1	17
function unknown (hypothetical protein)	904	8	2	1	33	28	27	5	61



Additional file 2: Diversity and distribution of 217 transcriptional regulators of P. aminophilus JCM 7686.

TABLE S2. Two-component systems and histidine kinases encoded by the *P. aminophilus* JCM 7686 genome.

Histidine kinase gene	Response regulator gene	Predicted function (based on the histidine kinase homology)	Replicon
JCM7686_pAMI4p337	JCM7686_pAMI4p336	regulation of a high affinity potassium-uptake	pAMI4
JCM7686_pAMI5p018	JCM7686_pAMI5p019	regulation of a trimethylamine N-oxide reductase respiratory system	pAMI5
JCM7686_pAMI5p061	JCM7686_pAMI5p060	unknown	pAMI5
JCM7686_pAMI5p082	JCM7686_pAMI5p081	unknown	pAMI5
JCM7686_pAMI5p117	JCM7686_pAMI5p118	regulation of expression of virulence factors	pAMI5
JCM7686_pAMI6p023	JCM7686_pAMI6p024	unknown	pAMI6
JCM7686_pAMI6p041	unknown	unknown	pAMI6
JCM7686_pAMI8p124	JCM7686_pAMI8p123	unknown	pAMI8
JCM7686_pAMI8p173	JCM7686_pAMI8p174	unknown	pAMI8
JCM7686_0157	JCM7686_0159	unknown	chromosome
JCM7686_0575	JCM7686_0576	regulation of nitrogen assimilation	chromosome
JCM7686_0577	JCM7686_0578	regulation of nitrogen assimilation	chromosome
JCM7686_0947	unknown	unknown	chromosome
JCM7686_1133	JCM7686_1132	unknown	chromosome
JCM7686_1281	JCM7686_1282	regulation of chemotaxis	chromosome
JCM7686_1314	unknown	unknown	chromosome
JCM7686_1371	JCM7686_1372	unknown	chromosome
JCM7686_1819	JCM7686_1818	unknown	chromosome
JCM7686_1927	JCM7686_1926	unknown	chromosome
JCM7686_2063	unknown	regulation of phosphate homeostasis	chromosome
JCM7686_2280	JCM7686_2279	unknown	chromosome
JCM7686_2539	unknown	regulation of differentiation and cell cycle progression	chromosome
JCM7686_2824	JCM7686_2823	regulation of C4-dicarboxylate metabolism	chromosome
JCM7686_2960	unknown	unknown	chromosome
JCM7686_3261	JCM7686_3260	unknown	chromosome
JCM7686_3369	JCM7686_3370	regulation of expression of virulence factors	chromosome
JCM7686_3383	JCM7686_3385	regulation of methanol and formaldehyde oxidation	chromosome
JCM7686_3423	JCM7686_3425	regulation of several anaerobic processes and assimilation of CO_2 and N_2	chromosome
JCM7686_3465	JCM7686_3468	unknown	chromosome

TABLE S3. Toxin-antitoxin systems encoded by the *P. aminophilus* JCM 7686 genome.

Toxin gene	Toxin family	Antitoxin gene	Antitoxin family	Replicon
JCM7686_pAMI2p005	RelE/ParE	JCM7686_pAMI2p006	HTH_XRE	pAMI2
JCM7686_pAMI3p001	RelE/ParE	JCM7686_pAMI3p002	ParD	pAMI3
JCM7686_pAMI4p367	Doc	JCM7686_pAMI4p368	Phd	pAMI4
JCM7686_pAMI5p1216	VapC	JCM7686_pAMI5p215	Phd	pAMI5
JCM7686_pAMI6p005	CcdB	JCM7686_pAMI6p006	CcdA	pAMI6
JCM7686_pAMI7p005	RelE/ParE	JCM7686_pAMI7p004	ParD	pAMI7
JCM7686_pAMI8p005	RelE/ParE	JCM7686_pAMI8p006	ParD	pAMI8
JCM7686_pAMI8p048	VapC	JCM7686_pAMI8p049	Phd	pAMI8
JCM7686_pAMI8p076	RelE/ParE	JCM7686_pAMI8p075	ParD	pAMI8
JCM7686_pAMI8p106	HipA	JCM7686_pAMI8p105	HipB	pAMI8
JCM7686_1676	HipA	JCM7686_1677	HipB	chromosome

TABLE S4. Chaperones and co-chaperonins encoded by the *P. aminophilus* JCM 7686 genome.

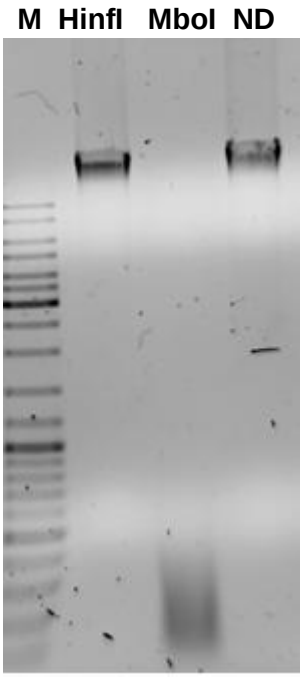
Chaperone gene	Homologous protein	Description	Replicon
JCM7686_pAMI4p204	GroEL/Hsp60	major molecular chaperone	pAMI4
JCM7686_pAMI5p111	HslJ	heat shock protein	pAMI5
JCM7686_pAMI5p166	DnaK/Hsp70	major molecular chaperone	pAMI5
JCM7686_pAMI6p154	ClpB-like	ATP-dependent Clp protease ATP-binding subunit	pAMI6
JCM7686_0077	GrpE	heat shock protein	chromosome
JCM7686_0087	ClpP	protease	chromosome
JCM7686_0138	ClpB	ATP-dependent Clp protease ATP-binding subunit	chromosome
JCM7686_0295	Lon	protease	chromosome
JCM7686_0479	FtsH	protease	chromosome
JCM7686_0482	HtpX	heat shock protein, Zn-dependent protease with chaperone function	chromosome
JCM7686_0582	Hfq	RNA chaperone	chromosome
JCM7686_0769	CspB	cold shock protein, RNA chaperone	chromosome
JCM7686_0997	CspA	cold shock protein, RNA chaperone	chromosome
JCM7686_1064	MoxR	MoxR-like ATPase	chromosome
JCM7686_1081	DnaJ-like	DnaJ-class molecular chaperone	chromosome
JCM7686_1391	DegP/HtrA	protease, protein quality control in the periplasm of Gram-negative bacteria	chromosome
JCM7686_1409	Lon	protease	chromosome
JCM7686_1579	ClpX	ATPase with chaperone activity	chromosome
JCM7686_1580	ClpP	protease	chromosome
JCM7686_1874	OmpH	outer membrane chaperone Skp (OmpH)	chromosome
JCM7686_1935	GroES/Hsp10	co-chaperonin	chromosome
JCM7686_1936	GroEL/Hsp60	major molecular chaperone	chromosome
JCM7686_1980	DnaJ-like	DnaJ-class molecular chaperone	chromosome
JCM7686_2155	ClpA	ATPase with chaperone activity	chromosome
JCM7686_2353	IbpA	molecular chaperone	chromosome
JCM7686_2458	CsaA	export-related chaperone	chromosome
JCM7686_2465	Hsp20	small heat shock protein	chromosome
JCM7686_2697	DnaJ-like	DnaJ-class molecular chaperone	chromosome
JCM7686_2977	HslO/Hsp33	disulfide bond chaperones of the HSP33 family	chromosome
JCM7686_3036	DnaK/Hsp70	major molecular chaperone	chromosome
JCM7686_3037	DnaJ	major molecular chaperone	chromosome
JCM7686_3176	ATP12	chaperone required for the assembly of the mitochondrial F1-ATPase	chromosome
JCM7686_3428	SecB	protein-export chaperone	chromosome
JCM7686_3434	HslU/ClpQ	ATP-dependent protease ATP-binding subunit	chromosome
JCM7686_3436	HslV/ClpY	ATP-dependent HslUV protease	chromosome

TABLE S5. Genes encoding DNA repair related proteins within the *P. aminophilus* JCM 7686 genome.

Gene	Genes in *E. coli**	Genes in *C. crescentus***	Description	Replicon
Base excision repair				
JCM7686_2055	*alkA*	CC_2201	3-methyl-adenine DNA glycosylase II	chromosome
JCM7686_0666	*fpg* (*mutM*)	CC_3707	formamidopyrimidine DNA glycosylase	chromosome
JCM7686_3073	*mutY*	CC_0377	adenine glycosylase	chromosome
JCM7686_0624	*nth*	CC_2272 CC_3731	endonuclease III; specific for apurinic and/or apyrimidinic sites	chromosome
JCM7686_0517	*tag*	CC_0382	3-methyl-adenine DNA glycosylase I	chromosome
JCM7686_0757 JCM7686_pAMI4p331	*ung* (*udg*)	CC_1333 CC_1549 CC_2084 CC_2333	uracil-DNA glycosylase	chromosome pAMI4
JCM7686_2891 JCM7686_3402	*xthA*	CC_2011 CC_3706	exodeoxyribonuclease III	chromosome
Nucleotide excision repair				
JCM7686_2156	*mfd*	CC_1844	transcription-repair coupling factor	chromosome
JCM7686_2515	*uvrA*	CC_2590	DNA damage recognition protein	chromosome
JCM7686_2310	*uvrB*	CC_2981	DNA damage binding protein	chromosome
JCM7686_3046	*uvrC*	CC_2881	excision nuclease	chromosome
JCM7686_0943	*uvrD*	CC_1528	DNA-dependent ATPase I and helicase II	chromosome
Mismatch repair				
JCM7686_3269	*mutL*	CC_0695	DNA mismatch repair protein	chromosome
JCM7686_3143	*mutS*	CC_0012	DNA mismatch repair protein	chromosome
Direct repair				
JCM7686_0623 JCM7686_pAMI5p076	*ada*	CC_0709 CC_3729	O6-methylguanine-DNA methyltransferase; transcription activator/repressor enzyme	chromosome pAMI5
JCM7686_pAMI5p075	*alkB*	CC_0009	1-methyladenine and 3-methylcytosine repair protein	pAMI5
JCM7686_2048	absent	CC_1330	DNA repair photolyase (SplB-like)	chromosome
JCM7686_2677	*ogt*	CC_0659 CC_0689	O-6-alkylguanine-DNA/cysteine-protein methyltransferase	chromosome
Recombinational repair				
JCM7686_3439	absent	CC_3538	ATP-dependent helicase/nuclease (AddA)	chromosome
JCM7686_3440	absent	CC_3537	ATP-dependent nuclease (AddB)	chromosome
JCM7686_2060	*radA*	CC_1660	DNA repair protein	chromosome
JCM7686_2538	*recA*	CC_1087	DNA strand exchange and renaturation, DNA-dependent ATPase, DNA- and ATP-dependent coprotease	chromosome
JCM7686_0660	*recF*	CC_0159	DNA replication and repair protein	chromosome
JCM7686_2217	*recG*	CC_1437	DNA helicase, resolution of Holliday junctions, branch migration	chromosome
JCM7686_1078	*recJ*	CC_1386	single-stranded-DNA-specific (5'→3') exonuclease	chromosome
JCM7686_0693	*recN*	CC_1983	recombination and repair protein	chromosome
JCM7686_0658	*recO*	CC_1564	recombinational repair protein	chromosome
JCM7686_0328	*recQ*	CC_3465	ATP-dependent DNA helicase	chromosome
JCM7686_1942	*recR*	CC_0269	recombination and repair protein	chromosome

JCM7686_0483	*ruvA*	CC_3237	Holliday junction helicase subunit A; branch migration	chromosome
JCM7686_0481	*ruvB*	CC_1283 CC_3236	Holliday junction helicase subunit B; branch migration	chromosome
JCM7686_0484	*ruvC*	CC_3238	Holliday junction nuclease; resolution of structures	chromosome
JCM7686_0280	*sbcB*	absent	exodeoxyribonuclease I	chromosome
JCM7686_2161 JCM7686_pAMI8p047	*ssb*	CC_1468	ssDNA-binding protein	Chromosome pAMI8
Other repair related proteins				
JCM7686_1991 JCM7686_2612 JCM7686_2906	*dinB* (*polIV*)	CC_2466	polymerase IV (damage-inducible protein B)	chromosome
JCM7686_1318 JCM7686_1972 JCM7686_pAMI8p132 JCM7686_pAMI8p203	*dnaE*	CC_1926 CC_3211	subunit alpha of DNA polymerase III	chromosome pAMI8
JCM7686_2872	*dut*	CC_3713	dUTPase	chromosome
JCM7686_0753	*lexA*	CC_1902	regulator for SOS regulon	chromosome
JCM7686_2218	*ligA*	CC_1522	DNA ligase NAD-dependent	chromosome
JCM7686_pAMI4p364	absent	CC_3610	DNA ligase ATP-dependent	pAMI4
JCM7686_pAMI4p365	*lhr*	CC_2040	Lhr-like helicases	pAMI4
JCM7686_0135 JCM7686_2983 JCM7686_2998 JCM7686_pAMI5p033	*mutT*	CC_0266 CC_0407 CC_0418 CC_0833 CC_1126 CC_1346 CC_1554 CC_2652 CC_3382 CC_3440 CC_3650	mutT/NUDIX family hydrolase	chromosome pAMI5
JCM7686_3038 JCM7686_pAMI8p117	*radC*	CC_2680 CC_2744	DNA repair protein	chromosome pAMI8
JCM7686_pAMI8p201	absent	CC_3213	ImuA repair protein	pAMI8
JCM7686_1973 JCM7686_pAMI8p202	absent	CC_3212	ImuB, Y-family DNA polymerase	chromosome pAMI8

* representative of *Gammaproteobacteria*
** representative of *Alphaproteobacteria*



HinfI – recognizes and cleaves the sequence 5'-GANTC-3'
MboI – recognizes and cleaves the sequence 5'-GATC-3'
For the analysis also other restriction endonucleases were used.

Inability to cleave GANTC sequences within *P. aminophilus* JCM 7686 genomic DNA shows that the sequences were methylated (and therefore protected) by CcrM methylase. Analogous analysis was performed with the cloned JCM7685_3079 gene in an *in vitro* experiment.

Additional file 7: Restriction patterns of *P. aminophilus* JCM 7686 genomic DNA cleaved with selected restriction endonucleases showing the protection of GANTC sites by the CcrM methylase (JCM7685_3079). ND - undigested DNA. M – GeneRuler 100–10,000 bp size marker.

TABLE S6. Summary of the sensitivity of various restriction endonucleases to DNA modifications introduced by the JCM7686_0772 and JCM7686_2655 proteins (m^5C MTases).

REase	Recognition site	Sensitivity to m^5C	Number of sites in pET28_JCM7686_0772	Number of sites in pET28_JCM7686_2655	Cleavage
AluI	AG**C**T	S	29	29	P, L
BglI	GCcNNNNNGG**C**	S	3	3	P, L
Bme1390I	CCNGG	U	32	31	P
BsuRI	GG**C**C	S	51	52	P, L
CaiI	CAGNNN**C**TG	S	4	4	P, L
EcoO109I	RGGN**CC**Y	S	4	5	P
EcoRII	C**C**WGG	S	12	12	P, L
Hin6I	G**C**GC	S	84	91	P
HpaII	**CC**GG	S	43	44	P
HpyF10VI	GCNNNNNNNG**C**	S	80	80	P
MspI	**C**cGG	S	43	44	P
NciI	C**C**SGG	S	20	19	P, L
PaeI	GCGCGC	U	8	10	P
PvuII	CAG**C**TG	S	4	6	P, L
TaiI	A**C**GT	S	16	17	P
XhoI	CTCGAG	U	3	2	P, L
Bsh1236I	**C**G**C**G	S	75	72	Y
Bsp68I	TCGCGA	U	3	3	Y
Bsp143I	GAT**C**	S	40	43	Y
BspLI	GGNNC**C**	S	31	32	Y
Eco130I	CCWWGG	U	6	6	Y
FspBI	CTAG	U	7	7	Y
Hin1II	CATG	U	41	41	Y
HinfI	GANT**C**	S	20	21	Y
MbiI	CCGCT**C**	S	7	6	Y
MluI	ACG**C**GT	S	1	1	Y
NcoI	**C**CATGG	S	3	2	Y
PamDI	**C**CATGG	S	3	2	Y

Y – complete cleavage; P – partial digestion; L – partial plasmid cleavage; S – sensitive to m^5C; U – unknown sensitivity to m^5C; **C** – the enzyme will not cleave if the **marked** cytosine is methylated; c – the enzyme is not sensitive to methylation of the marked cytosine; C – there are no data on the sensitivity of the enzyme to methylation of the unmarked cytosine.

TABLE S7. Summary of the sensitivity of various restriction endonucleases to DNA modifications introduced by the JCM7686_0815 protein (m^4C MTase).

REase	Recognition site	Sensitivity to m^4C	Number of sites in pET30_JCM7686_0815	Cleavage
CfrI	YGGCCR	U	5	P, L
HaeIII	GGC**C**	S	26	P*
Hin6I	GCGC	U	53	Y
HpaII	**CC**GG	S	34	Y
MspI	**C**cGG	S	34	Y

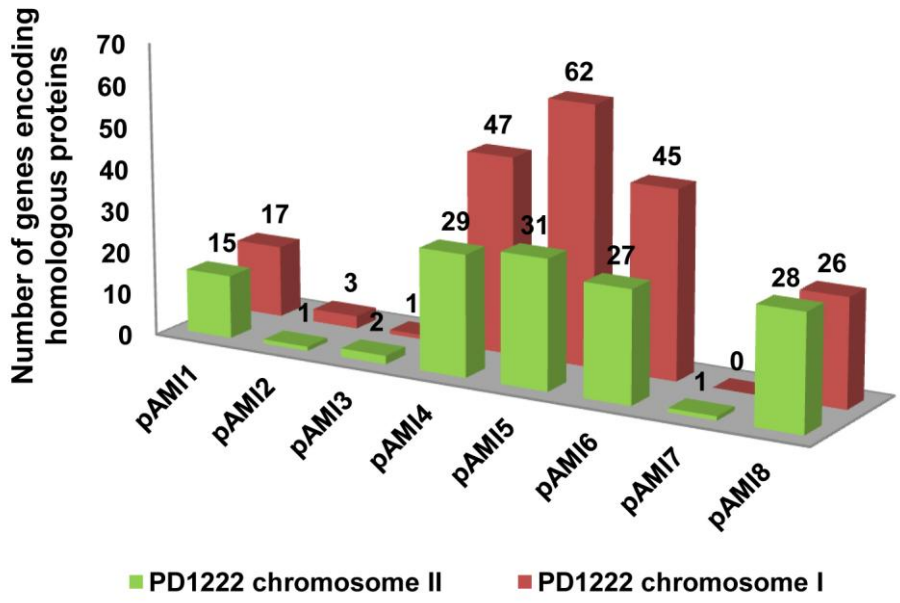
* the larger fused DNA fragments were observed only in case of incomplete cleavage of BsuRI sites that overlapped CfrI target sequences (Y**GGCC**R).

Y – complete cleavage; P – partial digestion; L – partial plasmid cleavage; S – sensitive to m^4C; U – unknown sensitivity to m^4C; **C** – the enzyme will not cleave if the **marked** cytosine is methylated; c – the enzyme is not sensitive to methylation of the marked cytosine; C – there are no data on the sensitivity of the enzyme to methylation of the unmarked cytosine.

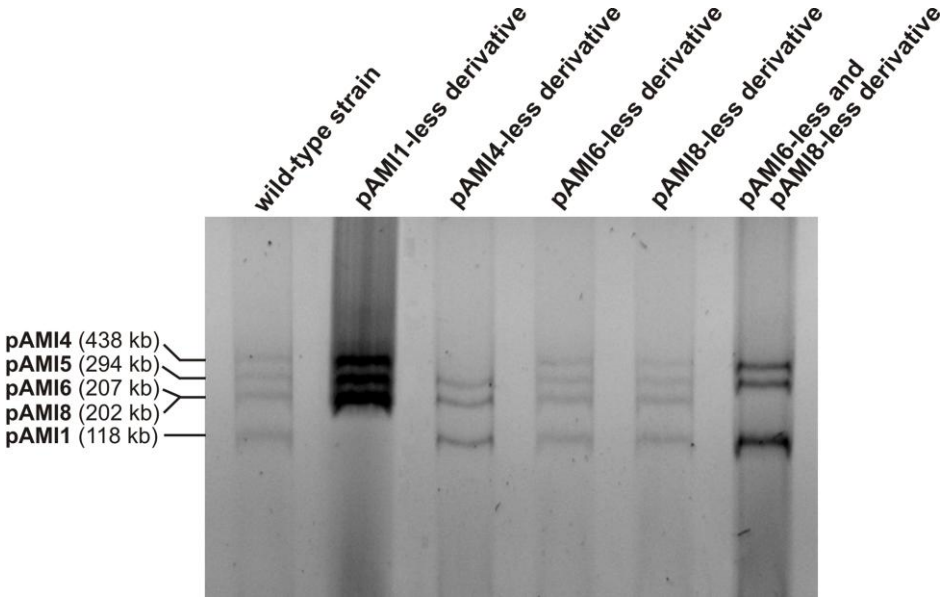
TABLE S8. Summary of the sensitivity of various restriction endonucleases to DNA modifications introduced by the JCM7686_1231, JCM7686_2255 and JCM7686_2934 proteins (m^6A MTases).

REase	Recognition site	Sensitivity to m^6A	Number of sites in pET28_JCM7686_2934	Number of sites in pET28_JCM7686_2255	Number of sites in pET28_JCM7686_1231	Cleavage
HinfI	G**A**NTC	S	19	19	18	N
AluI	**A**GCT	S	25	25	22	Y
CseI	G**A**CGC	S	14	14	13	Y
NlaIII	C**A**TG	U	32	32	30	Y
TaiI	**A**CGT	S	14	14	14	Y
TasI	**A**ATT	U	23	23	21	Y
Tru1I	TTA**A**	U	25	25	25	Y

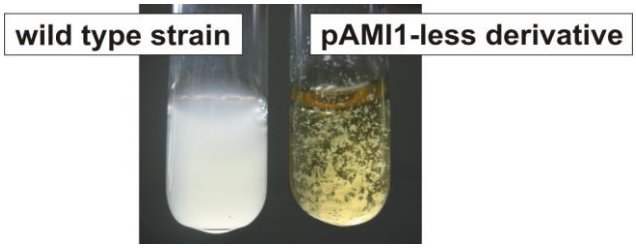
Y – complete cleavage; N – no digestion; S – sensitive to m^6A; U – unknown sensitivity to m^6A; **A** – the enzyme will not cleave if the **marked** adenine is methylated; A – there are no data on the sensitivity of the enzyme to methylation of the unmarked adenine.



Additional file 11: The number of homologous proteins encoded by JCM 7686 plasmids and *P. denitrificans* PD1222 chromosomes. The analysis was performed using the GeneOrder 4.0 tool. Proteins were considered homologous only if the BLAT threshold scores were >200. The results were verified manually by BLAST comparisons.



Additional file 12: Plasmid profiles of the JCM 7686 wild-type strain and its derivatives deprived of particular megasized-replicons.

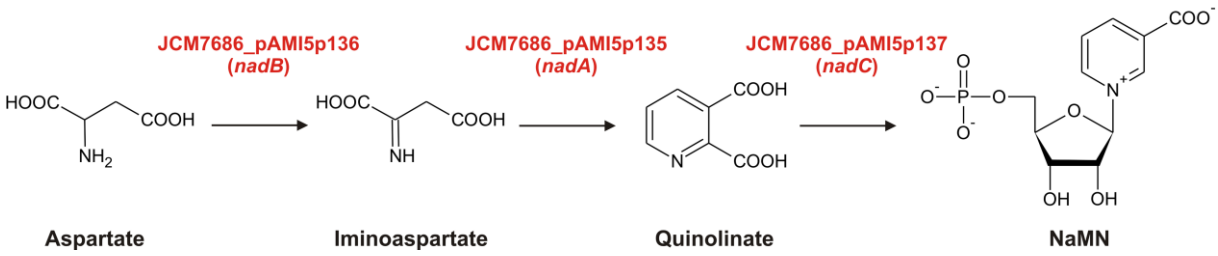


Additional file 13: The growth mode of wild-type JCM 7686 and the pAMI1-less derivative in liquid LB medium.

TABLE S9. Host ranges of *P. aminophilus* JCM 7686 pAMI1-8 replicons.

Host \ Plasmid	pAMI1	pAMI2	pAMI3	pAMI4	pAMI5	pAMI6	pAMI7	pAMI8
Alphaproteobacteria								
Agrobacterium tumefaciens LBA288	+	+	+	+	+	+	+	+
Ochrobactrum sp. LM19R	+	+	+	+	-	-	+	+
Paracoccus versutus UW225	+	+	+	+	+	+	+	+
Paracoccus pantotrophus KL100	+	+	+	+	+	+	+	+
Rhizobium etli CE3	+	+	+	+	+	+	+	+
Sinorhizobium sp. LM 21R	+	+	+	+	-	-	+	+
Betaproteobacteria								
Alcaligenes sp. LM16R	-	-	-	-	-	-	-	-
Gammaproteobacteria								
Escherichia coli BR825	-	-	-	-	-	-	-	-

- unable to replicate
+ able to replicate



Additional file 15: Schematic diagram of de novo NAD biosynthesis from aspartate to nicotinic acid mononucleotide (NaMN)..

Additional file 16: Transmission electron micrograph of tailed bacteriophage ϕPam-6 of *P. aminophilus* JCM 7686.

TABLE S10. Transposase genes within the *P. aminophilus* JCM 7686 genome.

Transposase gene	Truncated/complete	IS name/ Tn name	IS family (IS group)/ Tn family	IS length/ Tn length	Insertion site of complete TEs	Replicon
JCM7686_pAMI1p053	complete	Tn*3434a*	Tn*3*	3695	putrescine transport system permease (*potB*) gene	pAMI1
JCM7686_pAMI2p013	complete	IS*Pam3*	IS*3* (IS*407*)	1197	gene encoding hypothetical protein	pAMI2
JCM7686_pAMI2p014	complete					
JCM7686_pAMI2p018	complete	IS*Pam4*	IS*5* (IS*427*)	865	intergenic region	pAMI2
JCM7686_pAMI2p019	complete					
JCM7686_pAMI4p209	complete	IS*Pam1*	IS*5* (IS*903*)	1050	intergenic region	pAMI4
JCM7686_pAMI4p225	complete	IS*Pam5*	IS*3* (IS*407*)	1197	intergenic region	pAMI4
JCM7686_pAMI4p226	complete					
JCM7686_pAMI4p304	truncated	N/A	IS*1182*	N/A	N/A	pAMI4
JCM7686_pAMI4p305	truncated	N/A	IS*21*	N/A	N/A	pAMI4
JCM7686_pAMI4p306	complete	IS*Pam5*	IS*3* (IS*407*)	1197	intergenic region	pAMI4
JCM7686_pAMI4p307	complete					
JCM7686_pAMI4p333	truncated	N/A	IS*5* (IS*427*)	N/A	N/A	pAMI4
JCM7686_pAMI4p334	complete	IS*Pam5*	IS*3* (IS*407*)	1197	transposase gene	pAMI4
JCM7686_pAMI4p335	complete					
JCM7686_pAMI4p348	truncated	N/A	IS*3* (IS*150*)	N/A	N/A	pAMI4
JCM7686_pAMI4p349	truncated	N/A	IS*3* (IS*150*)	N/A	N/A	pAMI4
JCM7686_pAMI4p350	truncated	N/A	IS*3* (IS*150*)	N/A	N/A	pAMI4
JCM7686_pAMI4p353	truncated	N/A	IS*5* (IS*903*)	N/A	N/A	pAMI4
JCM7686_pAMI4p354	truncated	N/A	IS*5* (IS*903*)	N/A	N/A	pAMI4
JCM7686_pAMI4p357	truncated	N/A	IS*256*	N/A	N/A	pAMI4
JCM7686_pAMI4p358	truncated	N/A	IS*3* (IS*51*)	N/A	N/A	pAMI4
JCM7686_pAMI4p359	complete	IS*Pam5*	IS*3* (IS*407*)	1197	transposase gene	pAMI4
JCM7686_pAMI4p360	complete					
JCM7686_pAMI4p361	complete	IS*Pam6*	IS*1182*	1657	intergenic region	pAMI4
JCM7686_pAMI5p071	truncated	N/A	IS*5* (IS*903*)	N/A	N/A	pAMI5
JCM7686_pAMI5p092	complete	IS*Pam2*	IS*5* (IS*903*)	1054	ABC-type nitrate/sulfonate/bicarbonate transport system, ATPase gene	pAMI5
JCM7686_pAMI5p247	truncated	N/A	IS*3* (IS*407*)	N/A	N/A	pAMI5
JCM7686_pAMI5p248	truncated	N/A	IS*3* (IS*407*)	N/A	N/A	pAMI5
JCM7686_pAMI7p010	complete	Tn*3434a*	Tn*3*	3695	gene encoding hypothetical protein	pAMI7
JCM7686_pAMI8p011	complete	IS*Pam2*	IS*5* (IS*903*)	1054	DNA/RNA helicase (superfamily II, SNF2 family) gene	pAMI8
JCM7686_pAMI8p022	complete	IS*Pam7*	IS*3* (IS*407*)	1271	intergenic region	pAMI8
JCM7686_pAMI8p023						
JCM7686_pAMI8p024	truncated	N/A	IS*L3*	N/A	N/A	pAMI8

JCM7686_pAMI8p025 JCM7686_pAMI8p026	complete (frameshift)	IS*Pam8*	IS*66*	2669	transposase gene	pAMI8
JCM7686_pAMI8p090	complete	N/A	IS*3* (IS*150*)	N/A	N/A	pAMI8
JCM7686_pAMI8p091	truncated					
JCM7686_pAMI8p167	complete	Tn*3434a*	Tn*3*	3695	Na(+)-translocating NADH-quinone reductase subunit C gene	pAMI8
JCM7686_pAMI8p168	complete	IS*Pam1*	IS*5* (IS*903*)	1050	intergenic region	pAMI8
JCM7686_pAMI8p175	truncated	N/A	IS*21*	N/A	N/A	pAMI8
JCM7686_pAMI8p176	truncated	N/A	IS*21*	N/A	N/A	pAMI8
JCM7686_pAMI8p183	complete	Tn*3434a*	Tn*3*	3695	sodium/glutamate symporter gene	pAMI8
JCM7686_pAMI8p196	truncated	N/A	IS*481*	N/A	N/A	pAMI8
JCM7686_pAMI8p197	complete	IS*Pam9*	IS*110*	N/A*	transposase gene	pAMI8
JCM7686_pAMI8p198	truncated	N/A	IS*3* (IS*51*)	N/A	N/A	pAMI8
JCM7686_pAMI8p199	complete	IS*Pam10*	IS*5* (IS*903*)	1054	transposase gene	pAMI8
JCM7686_0276	complete	IS*Pam5*	IS*3* (IS*407*)	1197	transposase gene	chromosome
JCM7686_0277	complete					
JCM7686_0278	truncated	N/A	IS*3* (IS*407*)	N/A	N/A	chromosome
JCM7686_0279	truncated	N/A	IS*3* (IS*407*)	N/A	N/A	chromosome
JCM7686_0292	complete	IS*Pam5*	IS*3* (IS*407*)	1197	gene encoding hypothetical protein	chromosome
JCM7686_0293	complete					
JCM7686_1270	truncated	N/A	IS*3* (IS*51*)	N/A	N/A	chromosome
JCM7686_1274	truncated	N/A	IS*66*	N/A	N/A	chromosome
JCM7686_1275	truncated					
JCM7686_1679	truncated	N/A	IS*66*	N/A	N/A	chromosome
JCM7686_1868	truncated	N/A	IS*3* (IS*407*)	N/A	N/A	chromosome
JCM7686_1869	truncated					
JCM7686_2162	complete	IS*Pam5*	IS*3* (IS*407*)	1197	intergenic region	chromosome
JCM7686_2163	complete					
JCM7686_2224	truncated	N/A	IS*3* (IS*407*)	N/A	N/A	chromosome
JCM7686_2225	truncated					
JCM7686_2291	complete	N/A	Mu-type	N/A	N/A	chromosome
JCM7686_2292	complete					
JCM7686_2296	complete	IS*Pam11*	IS*3* (IS*51*)	1219	transposase gene	chromosome
JCM7686_2297	complete					
JCM7686_2298	truncated	N/A	IS*5* (IS*427*)	N/A	N/A	chromosome
JCM7686_2387	complete	IS*Pam5*	IS*3* (IS*407*)	1197	intergenic region	chromosome
JCM7686_2388	complete					
JCM7686_2390	truncated	N/A	IS*21*	N/A	N/A	chromosome
JCM7686_2391	truncated					

* Since ISs of the IS*110* family do not have IRs, and do not generate DRs, it is not possible to determine the length of IS*Pam9*.

Table S11. Bacterial strains used in this study.

Strain	Characteristics	Reference
Agrobacterium tumefaciens LBA288	Rif^r	[1]
Alcaligenes sp. LM16R	Rif^r derivative of wild-type strain LM16	[2]
Escherichia coli BR825	*polA::Tn10 trp*	[3]
E. coli ER2566	F$^-$ λ$^-$ *fhuA2* (*lon*) *ompT lacZ::T7 geneI gal sulA11 Δ(mcrC mrr)114::IS10 R*(*mcr-73*::*mini-Tn10)2 R*(*zgb-210::*Tn*10)1* (Tet^S) *endA1* (*Dcm*)	Laboratory collection
E. coli TG1	F' [*traD36 proAB*$^+$ *lacI*^{q} *lacZΔM15*]*supE thi-1 Δ(lac-proAB) Δ(mcrB-hsdSM)5, (*$r_K^-m_K^-$*)*	Laboratory collection
Ochrobactrum sp. LM19R	Rif^r derivative of wild-type strain LM19	[2]
Paracoccus aminophilus JCM 7686	wild type	[4]
P. aminophilus JCM 7686R	Rif^r derivative of wild-type strain JCM 7686	[5]
P. aminophilus AMI1	pAMI1-less derivative of the strain JCM 7686R	This work
P. aminophilus AMI4	pAMI4-less derivative of the strain JCM 7686R	This work
P. aminophilus AMI6	pAMI6-less derivative of the strain JCM 7686R	This work
P. aminophilus AMI8	pAMI8-less derivative of the strain JCM 7686R	This work
P. aminophilus AMI68	pAMI6-less and pAMI8-less derivative of the strain JCM 7686R	This work
Paracoccus pantotrophus KL100	Rif^r derivative of wild-type strain DSM 11073	[6]
Paracoccus versutus UW225	Rif^r derivative of wild-type strain UW1	[7]
Rhizobium etli CE3	Str^r	[8]
Sinorhizobium sp. LM 21R	Rif^r derivative of wild-type strain LM21	[2]

References

1. Hooykaas PJ, den Dulk-Ras H, Schilperoort RA: **Molecular mechanism of Ti plasmid mobilization by R plasmids: isolation of Ti plasmids with transposon-insertions in *Agrobacterium tumefaciens*.** *Plasmid* 1980, **4**:64-75.
2. Dziewit L, Adamczuk M, Szuplewska M, Bartosik D: **DIY series of genetic cassettes useful in construction of versatile vectors specific for Alphaproteobacteria**. *J Microbiol Methods* 2011, **86**(2):166-174.
3. Ludtke DN, Eichorn BG, Austin SJ: **Plasmid-partition functions of the P7 prophage**. *J Mol Biol* 1989, **209**(3):393-406.
4. Urakami T, Araki H, Oyanagi H, Suzuki K, Komagata K: ***Paracoccus aminophilus* sp. nov. and *Paracoccus aminovorans* sp. nov., which utilize *N,N*-dimethylformamide**. *Int J Syst Bacteriol* 1990, **40**(3):287-291.
5. Dziewit L, Dmowski M, Baj J, Bartosik D: **Plasmid pAMI2 of *Paracoccus aminophilus* JCM 7686 carries *N,N*-dimethylformamide degradation-related genes whose expression is activated by a LuxR family regulator**. *Appl Environ Microbiol* 2010, **76**(6):1861-1869.
6. Bartosik D, Baj J, Bartosik AA, Wlodarczyk M: **Characterization of the replicator region of megaplasmid pTAV3 of *Paracoccus versutus* and search for plasmid-encoded traits**. *Microbiology (Reading, England)* 2002, **148**(Pt 3):871-881.
7. Bartosik D, Baj J, Plasota M, Piechucka E, Wlodarczyk M: **Analysis of *Thiobacillus versutus* pTAV1 plasmid functions.** *Acta Microbiol Pol* 1993, **39**:5-11.
8. Noel KD, Sanchez A, Fernandez L, Leemans J, Cevallos MA: ***Rhizobium phaseoli* symbiotic mutants with transposon Tn5 insertions**. *J Bacteriol* 1984, **158**(1):148-155.

Table S12. Plasmids used and constructed in this study.

Plasmid	Characteristics	Reference
pABW1	Km^r; *ori* pMB1; mobilizable cloning vector; *oriT* RK2	[1]
pABW1-AMI1	Km^r; *ori* pMB1; pABW1 derivative carrying 6.8-kb BamHI restriction fragment of pAMI1 (contains REP and PAR modules)	This work
pABW1-AMI4	Km^r; *ori* pMB1; pABW1 derivative carrying 5.6-kb SalI/KpnI restriction fragment of pAMI4 (contains REP module)	This work
pABW1-AMI5	Km^r; *ori* pMB1; pABW1 derivative carrying 4.0-kb DNA region (contains REP and PAR modules) of pAMI5 (amplified by PCR with primers LAMI5 and RAMI5) inserted between XbaI and SphI sites	This work
pABW1-AMI6	Km^r; *ori* pMB1; pABW1 derivative carrying 1.9-kb DNA region (contains REP module) of pAMI6 (amplified by PCR with primers LAMI6 and RAMI62) inserted between XbaI and SphI sites	This work
pABW1-AMI8	Km^r; *ori* pMB1; pABW1 derivative carrying 4.3-kb DNA region (contains REP and PAR modules) of pAMI8 (amplified by PCR with primers LAMI8 and RAMI8) inserted into EcoRI site	This work
pBBR1MCS-2	Km^r; *ori* pBBR1; broad-host-range cloning vector; mobilizable cloning vector, *oriT* RK2	[2]
pBBR-CcrM	Km^r; *ori* pBBR1; pBBR1MCS-2 derivative carrying 1.4-kb DNA region (contains JCM7686_3079 gene) amplified by PCR with primers FCCRMXB and RCCRMXH and inserted between XbaI and XhoI sites	This work
pET28_JCM7686_1231	Km^r; *ori* pMB1; pET-28a derivative carrying 0.6-kb DNA region (contains JCM7686_1231 gene) amplified by PCR with primers JCM7686_1231F and JCM7686_1231R and inserted between NdeI and SalI sites	This work
pET28_JCM7686_2255	Km^r; *ori* pMB1; pET-28a derivative carrying 0.6-kb DNA region (contains JCM7686_2255 gene) amplified by PCR with primers JCM7686_2255F and JCM7686_2255R	This work
pET28_JCM7686_2655	Km^r; *ori* pMB1; pET-28a derivative carrying 2.1-kb DNA region (contains JCM7686_2655 gene) amplified by PCR with primers JCM7686_2655F and JCM7686_2655R and inserted between NheI and HindIII sites	This work
pET28_JCM7686_2934	Km^r; *ori* pMB1; pET-28a derivative carrying 0.6-kb DNA region (contains JCM7686_2934 gene) amplified by PCR with primers JCM7686_2934F and JCM7686_2934R and inserted between NcoI and SalI sites	This work
pET28_JCM7686_0772	Km^r; *ori* pMB1; pET-28a derivative carrying 2.1-kb DNA region (contains JCM7686_0772 gene) amplified by PCR with primers JCM7686_0772F and JCM7686_0772R and inserted between NheI and HindIII sites	This work
pET-28a	Km^r; *ori* pMB1; expression vector with T7lac promoter	Novagen
pET30_JCM7686_3079	Km^r; *ori* pMB1; pET-30a derivative carrying 1.2-kb DNA region (contains JCM7686_3079 gene) amplified by PCR with primers JCM7686_3079F and JCM7686_3079R and inserted between NdeI and XhoI sites	This work
pET30_JCM7686_0815	Km^r; *ori* pMB1; pET-30a derivative carrying 0.5-kb DNA region (contains JCM7686_0815 gene) amplified by PCR with primers JCM7686_0815F and JCM7686_0815R and inserted between NdeI and XhoI sites	This work
pET-30a	Km^r; *ori* pMB1; expression vector with T7lac promoter	Novagen
pRK2013	Km^r; helper plasmid carrying genes for conjugal transfer of RK2	[3]

References

1. Bartosik D, Bialkowska A, Baj J, Wlodarczyk M: **Construction of mobilizable cloning vectors derived from pBGS18 and their application for analysis of replicator region of a pTAV202 mini-derivative of *Paracoccus versutus* pTAV1 plasmid**. *Acta Microbiol Pol* 1997, **46**(4):387-392.
2. Kovach ME, Phillips RW, Elzer PH, Roop RM, 2nd, Peterson KM: **pBBR1MCS: a broad-host-range cloning vector**. *BioTechniques* 1994, **16**(5):800-802.
3. Ditta G, Stanfield S, Corbin D, Helinski DR: **Broad host range DNA cloning system for gram-negative bacteria: construction of a gene bank of Rhizobium meliloti**. *Proc Natl Acad Sci USA* 1980, **77**(12):7347-7351.

Table S13. Oligonucleotide primers used in this study.

Primer	Sequence (5'->3')	Function
JCM7686_0772F	GTTGCTAGCATGACCTATCAGCCGCACCTGCC	amplification of JCM7686_0772
JCM7686_0772R	GTTAAGCTTTCACCCAACCCTCCTGTCTTCCTTGATG	amplification of JCM7686_0772
JCM7686_1231F	GTTCATATGGGTAAGCGATCGAATTTCC	amplification of JCM7686_1231
JCM7686_1231R	GTTTCAGTCGACTGCCGCTGCCCTCCCGAAG	amplification of JCM7686_1231
JCM7686_2255F	CAGCGTCGTATTGGCGGGCTATC	amplification of JCM7686_2255
JCM7686_2255R	GACCGAGAGCGGAATCGCTCATG	amplification of JCM7686_2255
JCM7686_2655F	GTTGCTAGCATGACCTATCAACCGCACCTGCC	amplification of JCM7686_2655
JCM7686_2655R	GTTAAGCTTTCACCCAACCCTCCTGTTTTCCTC	amplification of JCM7686_2655
JCM7686_2934F	GTTCCATGGGTAAGCGATCGAATTTCC	amplification of JCM7686_2934
JCM7686_2934R	GTTGTCGACTGCCGCGGCCCTCCCGATAAATTC	amplification of JCM7686_2934
JCM7686_0815F	GTTCATATGAGCGAAACTCCTGTGCTCGATG	amplification of JCM7686_0815
JCM7686_0815R	GTTTTACTCGAGTCGGGTGGCCTCCGGCTTCATG	amplification of JCM7686_0815
JCM7686_3079F	GTTCAT ATGAGGCAGAGCAAAATGAAAACA	amplification of JCM7686_3079
JCM7686_3079R	GTTTCACTCGAG GTTCGGGCGCTGTGCGCCGTCG	amplification of JCM7686_3079
FCCRMXB	GATCTAGAAAGTGTTGCGCGACAGACTC	amplification of JCM7686_3079
RCCRMXH	TACTCGAGCGGCGAAACGAATGGATCAG	amplification of JCM7686_3079
LAMI5	GCTCTAGATGGCGCAAGAAGATCGAAGG	amplification of a replication module of pAMI5
RAMI5	TAGCATGCGCGAAGCGGATTAACCATCG	amplification of a replication module of pAMI5
LAMI6	GACTCTAGAATCCGCTTCGGTCTTCTGG	amplification of a replication module of pAMI6
RAMI62	TAGCATGCGCGTTCTGCGAGATCCTGTC	amplification of a replication module of pAMI6
LAMI8	AGACCTGCAGGACTTCCACTGAATAGCAG	amplification of a replication module of pAMI8
RAMI8	CGAGAATCCACAGAAGCAGCTTGGTCATG	amplification of a replication module of pAMI8

Publikacja 3

Łukasz Dziewit*, **Jakub Czarnecki***, Emilia Prochwicz, Daniel Wibberg, Andreas Schlüter, Alfred Pühler, Dariusz Bartosik

Genome-guided insight into the methylotrophy of *Paracoccus aminophilus* JCM 7686

Frontiers in Microbiology, 2015, 6:852

Impact Factor: 3,989
Punktacja MNiSW: 35

frontiers in Microbiology

ORIGINAL RESEARCH
published: 21 August 2015
doi: 10.3389/fmicb.2015.00852

Genome-guided insight into the methylotrophy of *Paracoccus aminophilus* JCM 7686

Lukasz Dziewit[1*†], *Jakub Czarnecki*[1†], *Emilia Prochwicz*[1], *Daniel Wibberg*[2], *Andreas Schlüter*[2], *Alfred Pühler*[2] *and Dariusz Bartosik*[1]

[1] *Department of Bacterial Genetics, Institute of Microbiology, Faculty of Biology, University of Warsaw, Warsaw, Poland,*
[2] *Institute for Genome Research and Systems Biology, Center for Biotechnology (CeBiTec), Bielefeld University, Bielefeld, Germany*

Paracoccus aminophilus JCM 7686 (*Alphaproteobacteria*) is a facultative, heterotrophic methylotroph capable of utilizing a wide range of C1 compounds as sole carbon and energy sources. Analysis of the JCM 7686 genome revealed the presence of genes involved in the oxidation of methanol, methylamine, dimethylamine, trimethylamine, *N*,*N*-dimethylformamide, and formamide, as well as the serine cycle, which appears to be the only C1 assimilatory pathway in this strain. Many of these genes are located in different extrachromosomal replicons and are not present in the genomes of most members of the genus *Paracoccus*, which strongly suggests that they have been horizontally acquired. When compared with *Paracoccus denitrificans* Pd1222 (type strain of the genus *Paracoccus*), *P. aminophilus* JCM 7686 has many additional methylotrophic capabilities (oxidation of dimethylamine, trimethylamine, *N,N*-dimethylformamide, the serine cycle), which are determined by the presence of three separate gene clusters. Interestingly, related clusters form compact methylotrophy islands within the genomes of *Paracoccus* sp. N5 and many marine bacteria of the *Roseobacter* clade.

Keywords: *Paracoccus aminophilus* JCM 7686, methylotrophy, serine cycle, methanol, methylated amine, *N,N*-dimethylformamide, chromid, plasmid

OPEN ACCESS

Edited by:
Weiwen Zhang,
Tianjin University, China

Reviewed by:
Ludmila Chistoserdova,
University of Washington, USA
Ulrike Kappler,
University of Queensland, Australia

***Correspondence:**
Lukasz Dziewit,
Department of Bacterial Genetics, Institute of Microbiology, Faculty of Biology, University of Warsaw, Miecznikowa 1, 02-096 Warsaw, Poland
ldziewit@biol.uw.edu.pl

[†] *These authors have contributed equally to this work.*

Specialty section:
This article was submitted to Microbial Physiology and Metabolism, a section of the journal Frontiers in Microbiology

Received: *17 April 2015*
Accepted: *04 August 2015*
Published: *21 August 2015*

Citation:
Dziewit L, Czarnecki J, Prochwicz E, Wibberg D, Schlüter A, Pühler A and Bartosik D (2015) Genome-guided insight into the methylotrophy of Paracoccus aminophilus JCM 7686. Front. Microbiol. 6:852. doi: 10.3389/fmicb.2015.00852

Introduction

The genus *Paracoccus* (*Alphaproteobacteria*) comprises bacteria isolated from various pristine and polluted environments (e.g., soil, marine sediments, seawater, biofilters, activated sludge, or human tissues) (Urakami et al., 1990; Siller et al., 1996; Lipski et al., 1998; Tsubokura et al., 1999; Funke et al., 2004; Lee et al., 2004; Liu et al., 2006). These bacteria show diverse metabolic properties and can switch between different growth modes—for example, heterotrophic growth on a wide range of organic compounds vs. chemolithoautotrophic growth on reduced sulfur compounds, hydrogen or ferrous ions as the energy sources, or aerobic respiration vs. anaerobic nitrate respiration (Kelly et al., 2006). Because of their degradative capabilities, many *Paracoccus* spp. strains are suitable for application in bioremediation systems. They have been successfully employed in the bioremediation of soils contaminated with polycyclic aromatic hydrocarbons (PAHs, Sun et al., 2013) and in the removal of nitrate and *N,N*-dimethylformamide (DMF) from wastewater (Liu et al., 2012; Sanjeevkumar et al., 2013). Moreover, their ability to utilize pesticides and insecticides, including the highly toxic chlorpyrifos, 3,5,6-trichloro-2-pyridinol, methyl parathion, and carbonfuran, has also been demonstrated (Li et al., 2011).

About 50% of known *Paracoccus* spp. strains are described as methylotrophs, i.e., organisms utilizing C1 compounds (reduced carbon compounds containing no carbon-carbon bonds) as sole carbon and energy sources (Baker et al., 1998). Methylotrophs play an important role in global carbon, nitrogen, and sulfur cycling, and for this reason their biochemistry has been subjected to extensive studies (Trotsenko and Murrell, 2008; Chistoserdova, 2011). As described by Chistoserdova (2011), the methylotrophy process can be divided into three stages: (i) primary oxidation of C1 substrates, which results in formaldehyde, methyl- or methylene-tetrahydrofolate (CH_3–THF or CH_2=THF) formation, (ii) oxidation of formaldehyde, CH_3–THF or CH_2=THF to CO_2, and (iii) assimilation of C1 units. The final stage may be performed via the ribulose monophosphate (RuMP) cycle, the serine cycle or the Calvin–Benson–Bassham (CBB) cycle (Chistoserdova, 2011).

Paracoccus denitrificans Pd1222 (type strain of the genus *Paracoccus,* Kelly et al., 2006) exemplifies so-called autotrophic methylotrophs (Chistoserdova, 2011), assimilating CO_2 derived from the oxidation of C1 compounds [in Pd1222 these are methanol (MeOH) or methylamine (MA)] via the Calvin cycle (Baker et al., 1998). Other methylotrophic strains of *Paracoccus* spp. have been poorly characterized. Interestingly, in comparison with *P. denitrificans*, these strains show significant differences in their methylotrophic metabolism, not only in the range of C1 compounds utilized [many strains are able to oxidize trimethylamine (TMA), trimethylamine *N*-oxide (TMAO), dimethylamine (DMA), dichloromethane or DMF] (Urakami et al., 1990; Kim et al., 2001, 2003; Turova et al., 2001), but also in the central metabolic pathways mediating C1 unit assimilation (Beck et al., 2015).

Very recently, a set of genes encoding the enzymes of the serine cycle have been identified in *Paracoccus* sp. N5 (Beck et al., 2015). Since this strain also encodes all enzymes of the Calvin cycle it was classified as a facultatively autotrophic methylotroph (Beck et al., 2015). In this work we have characterized *Paracoccus aminophilus* JCM 7686, which represents another metabolic group of methylotrophic bacteria—the heterotrophic serine cycle methylotrophs. This strain does not encode ribulose-1,5-bisphosphate carboxylase/oxygenase (RuBisCO), crucial for the Calvin cycle, so its methylotrophic properties exclusively rely on the serine cycle.

The multipartite genome of *P. aminophilus* JCM 7686 was described in detail in our previous report (Dziewit et al., 2014). It is composed of a single circular chromosome (3.6 Mb) and eight circular extrachromosomal replicons. The functional analyses of those replicons revealed that six of them (pAMI1, pAMI2, pAMI3, pAMI4, pAMI7, and pAMI8) are plasmids and two (pAMI5 and pAMI6) are chromids (i.e., elements essential for host viability and sharing characteristics of both chromosomes and plasmids) (Dziewit et al., 2007, 2011a,b, 2014). In the present study, bioinformatic sequence analyses of the JCM 7686 genome together with functional characterization of selected genes have provided deeper insight into both the biochemistry of the methylotrophy of this strain and the role of the extrachromosomal genetic elements in determination of its methylotrophic capability.

Materials and Methods

Strains, Plasmids, and Culture Conditions

The strains used in this study are described in **Table S1**. All strains were grown in lysogeny broth (LB) (Sambrook and Russell, 2001) at 37°C [*E. coli* TG1 (Gibson, 1984) and S17-1 (Simon et al., 1983)] and 30°C [*P. aminophilus* JCM 7686R (Bartosik et al., 2002)]. *P. aminophilus* was also grown in minimal salts medium (AC) (Wood and Kelly, 1977) at 30°C. When necessary, the media were supplemented with kanamycin (50 μg/ml), tetracycline (2 μg/ml for *Paracoccus* spp. or 20 μg/ml for *E. coli*), chloramphenicol (12.5 μg/ml), rifampicin (50 μg/ml), gentamicin (10 μg/ml), or sucrose (11% w/v). The following compounds were used as the carbon source in minimal media: L-arabinose (0.2% w/v), *N,N*-dimethylformamide (20 mM), trimethylamine (20 mM), dimethylamine (10 mM), methylamine (10 mM), formamide (20 mM), and methanol (20 mM). All plasmids used [i.e., pBBR1MCS-3 (Kovach et al., 1994), pBBR1MCS-5 (Kovach et al., 1994), pDIY-KM (Dziewit et al., 2011a), pDS132 (Philippe et al., 2004), pKRP12 (Reece and Phillips, 1995), pRK2013 (Ditta et al., 1980)] and constructed in this study are described in **Table S2**.

Standard Genetic Manipulations

The isolation of DNA and common DNA manipulation methods were performed as described by Sambrook and Russell (2001). PCR was performed in an Eppendorf Mastercycler with pDIY-KM, pKRP12 or total DNA of *P. aminophilus* as the template, appropriate oligonucleotide primers (**Table S3**), dNTP mixture and Phusion High-Fidelity DNA polymerase (Thermo Scientific; with supplied HF buffer). Transformation of *E. coli* strains was performed according to the method of Kushner (1978). Bi- and tri-parental matings were performed as previously described (Bartosik et al., 2001). For triparental mating, overnight cultures of the donor strain *E. coli* TG1, the appropriate recipient strain, and *E. coli* DH5α carrying the helper plasmid pRK2013 were mixed in a ratio 1:2:1. For biparental mating, overnight cultures of the donor strain *E. coli* S17-1 carrying a mobilizable vector and the appropriate recipient strain were mixed in a ratio 1:1. Then 100 μl of the mixture was spread on a plate with solidified LB medium and incubated overnight at 30°C. Then the bacteria were washed off the plate, and the suitable dilutions were plated on appropriate selective media to select transconjugants (Bartosik et al., 2001).

Gene Disruption

Deletion of selected *P. aminophilus* JCM 7686 genes was performed using a gene replacement method. DNA cassettes for gene disruption, containing an antibiotic resistance gene flanked by PCR-amplified DNA fragments (ca. 500 bp) homologous to the DNA regions surrounding the gene targeted for disruption, were created by overlap extension PCR with specific primers (**Table S3**) or by restriction cloning (**Table S2**). The obtained DNA cassettes were transferred by biparental mating from *E. coli* S17-1 into rifampicin resistant *P. aminophilus* JCM 7686R cells on *sacB* gene-containing vector pDS132, unable to replicate in *Alphaproteobacteria*. Double-cross recombinants were selected

on appropriate medium containing rifampicin, sucrose and kanamycin or tetracycline, depending on the antibiotic resistance gene used for cassette construction. The correctness of the generated disruptions was verified by sequencing DNA fragments amplified by PCR using appropriate primer pairs (**Table S3**).

RT-qPCR

Total RNAs used for the RT-qPCR analyses were isolated from *P. aminophilus* cells grown on TMA or arabinose as the sole carbon and energy source. The strain was cultured in appropriate minimal media in three biological repeats and cells were collected during late exponential phase. RNA was isolated using TRI Reagent® Solution (Ambion) according to the manufacturer's recommendations. Contaminating DNA was removed using DNA-free™, DNase Treatment & Removal (Ambion) according to the manufacturer's recommendations. The total RNA was transcribed into cDNA using the Maxima First Strand cDNA Synthesis Kit for RT-qPCR (Thermo Scientific) according to the manufacturer's instruction. RT-qPCR reactions were carried out in LightCycler® 480 Instrument II (Roche) using 5 × HOT FIREPol® EvaGreen® qPCR Mix Plus (no ROX). Oligonucleotide primers used in the study are listed in **Table S3**. Relative quantification of gene transcription was performed using the comparative C_t (threshold cycle) method.

Bioinformatics

The putative function of particular genes was assigned using BLAST programs (Altschul et al., 1997) and PRIAM tool (Claudel-Renard et al., 2003) as previously described (Heinl et al., 2012). Metabolic pathways were recognized and described using Pathway tools (Karp et al., 2010), the MetaCyc database (Caspi et al., 2008) and GenDB 2.0 (Meyer et al., 2003). For the identification of methylotrophy-linked genes/proteins a BLASTn/BLASTp analysis comparing the sequence of each gene/protein with the genome/proteome of a particular *Paracoccus* strain was performed. Strict cutoff values were applied for this analysis: $e < 10^{-40}$, 70% for minimal query coverage and sequence identity of at least 45%.

Results and Discussion

Identification of *P. aminophilus* JCM 7686 Genes Linked to Methylotrophy

P. aminophilus JCM 7686 was isolated in Japan from soil contaminated with *N,N*-dimethylformamide (DMF) as a strain able to utilize many C1 compounds (**Figure S1**) (Urakami et al., 1990). However, no further analysis regarding the methylotrophy of this strain was performed. In the initial stage of this study we confirmed that, besides DMF, JCM 7686 can utilize methylamine, dimethylamine, trimethylamine, and formamide. We also found that it is able to utilize methanol (**Figure 1**), which is contrary to the original observations of Urakami et al. (1990). To determine the genetic basis of these phenotypes we examined the JCM 7686 genome (Dziewit et al., 2014) for the presence of genes linked to C1 metabolism and we performed functional analysis of selected genes to confirm our predictions. The collected data permitted reconstruction of the complex C1 metabolic pathway of this strain (**Figure 2**, **Table S4**).

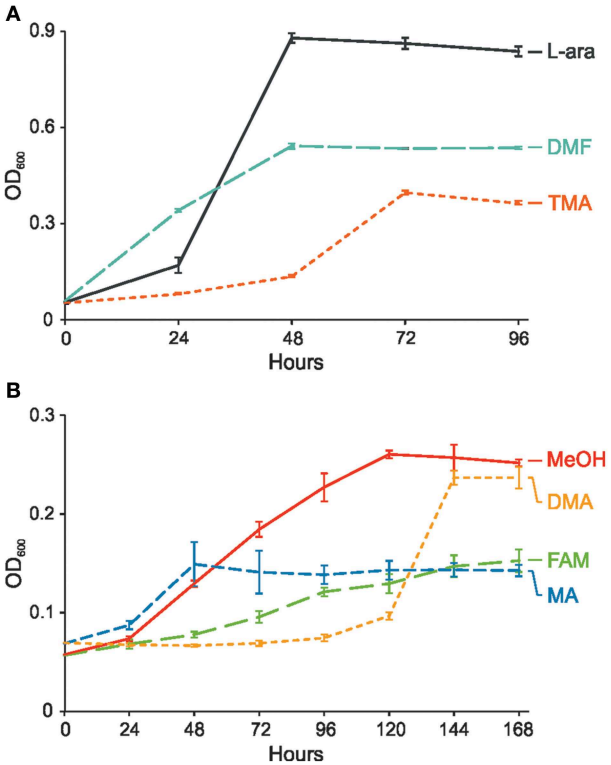


FIGURE 1 | Growth of *P. aminophilus* JCM 7686 on various C1 compounds or L-arabinose as a sole carbon and energy source. (A) Growth on L-arabinose (L-ara), *N,N*-dimethylformamide (DMF), and trimethylamine (TMA), **(B)** Growth on methanol (MeOH), dimethylamine (DMA), formamide (FAM), and methylamine (MA). The values are means of three replicates, and the error bars indicate the standard deviations.

Genes Involved in Methanol Utilization

The ability of *P. denitrificans* Pd1222 to utilize methanol was shown to be dependent on the presence of *mxa* genes encoding i.a. subunits of a heterotetrameric PQQ-dependent calcium-binding MeOH dehydrogenase (MxaFI) (Van Spanning et al., 1991). Although *P. aminophilus* JCM 7686 is also able to oxidize MeOH, the *mxa* genes were not detected in this strain. Nevertheless, a gene encoding another type of MeOH dehydrogenase (*xoxF*; JCM7686_0090) was identified within the JCM 7686 chromosome. XoxF represents a group of homodimeric methanol dehydrogenases, related to the large subunit of MxaFI, which bind rare-earth elements instead of calcium (Keltjens et al., 2014). Interestingly, *xoxF* genes are widespread among both methylotrophs and non-methylotrophic bacteria (Keltjens et al., 2014). Studies on *Rhodobacter sphaeroides* revealed that XoxF is required for methanol oxidation during both aerobic and anaerobic photosynthetic growth (Wilson et al., 2008).

The *P. aminophilus*-encoded XoxF belongs to the XoxF5 protein family (Keltjens et al., 2014). In the JCM 7686 chromosome, the *xoxF* gene is clustered together with *xoxG*

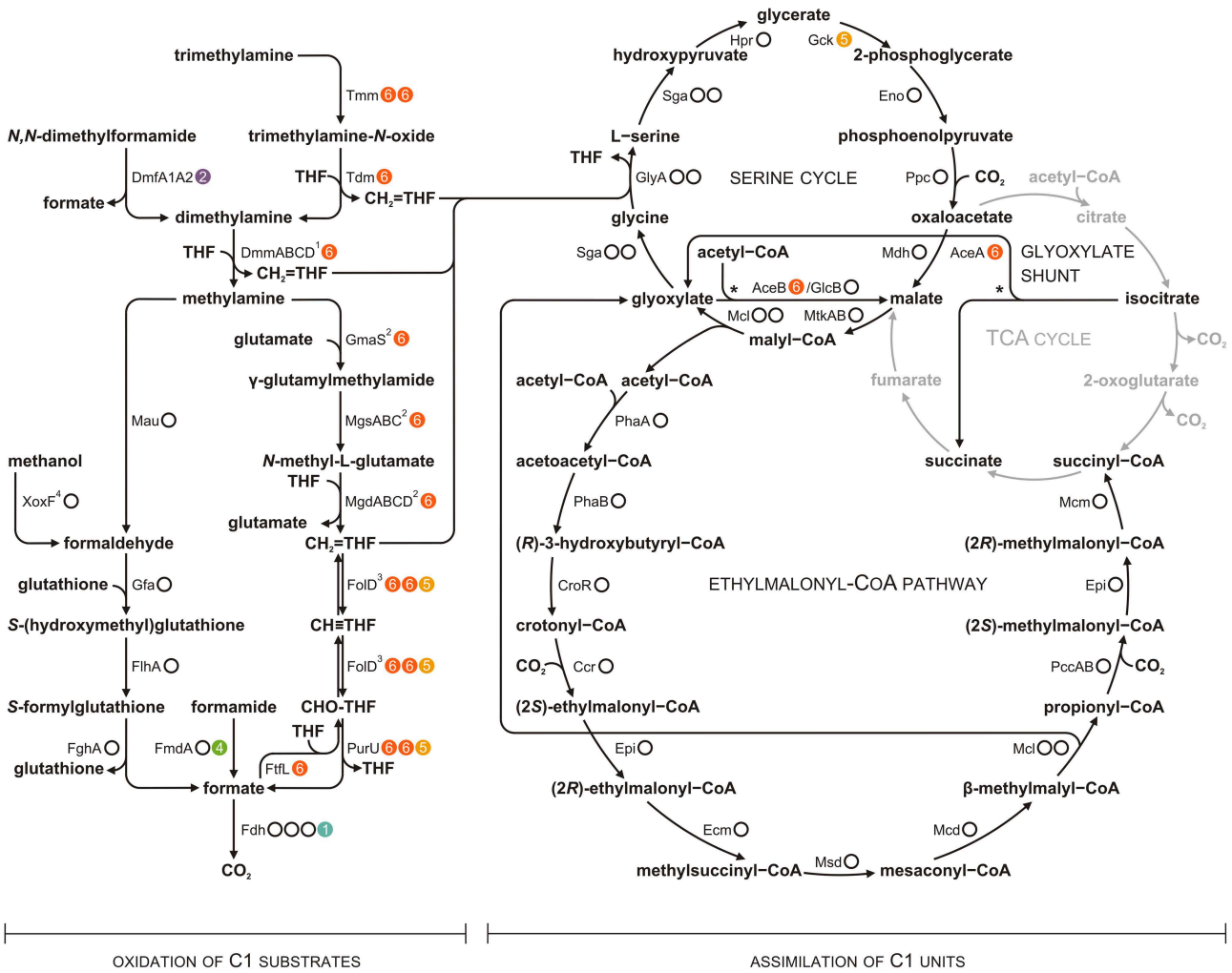


FIGURE 2 | Methylotrophic metabolism of *P. aminophilus* JCM 7686 reconstructed on the basis of genomic analysis. Dots to the right of the enzyme designations indicate the number of loci encoding the corresponding protein and their location in the genome (white dots, the chromosome; colored dots with numbers, extrachromosomal replicons pAMI1, pAMI2, pAMI4, pAMI5 or pAMI6). Full names of enzymes are given in **Table S4**. The glyoxylate shunt reactions are indicated by asterisks. [1]DmmABCD is proposed to be involved in DMA metabolism on the basis of mutational analysis (Beck et al., 2015 and this study). [2]The presented *N*-methylglutamate pathway (NMGP) was selected from among the possible schemes. Not only CH_2=THF, but also free formaldehyde is considered to be a product of this pathway (Nayak and Marx, 2014). [3]Previously, FolD was thought to act only in the oxidative direction (Chistoserdova, 2011), but it has recently been proposed to act bidirectionally, at least in some bacteria (Beck et al., 2015). [4]XoxF-type dehydrogenases may convert methanol directly into formate (Keltjens et al., 2014).

(JCM7686_0091, encoding a cytochrome c used as an electron acceptor for methanol oxidation) and *xoxJ* (JCM7686_0092, encoding a putative periplasmic binding protein), as well as with the genes of a glutathione-dependent formaldehyde dehydrogenase system (JCM7686_0085, JCM7686_0086, and JCM7686_0089) (**Table S4**). It suggests that formaldehyde, and not formate, is the product of methanol oxidation catalyzed by XoxF5 proteins (Keltjens et al., 2014). The clustering of *xoxF* and the genes of the glutathione-dependent formaldehyde oxidation pathway is also seen in many other alphaproteobacterial genomes, e.g., in *P. denitrificans* Pd1222, *Roseobacter litoralis* Och 149, *Dinoroseobacter shibae* DFL 12 and *Sinorhizobium fredii* HH103 (Keltjens et al., 2014 and this study).

To evaluate the role of the identified *xoxF* gene in methanol utilization, its mutational analysis was performed. This revealed that a strain with disrupted *xoxF* was unable to grow in minimal medium with methanol as the sole source of carbon. The wild-type phenotype was restored when the *xoxF* gene cloned in vector pBBR1MCS-3 was introduced into the mutant cells (**Figure 3**). This confirmed that XoxF is the enzyme responsible for methanol oxidation in *P. aminophilus* JCM 7686.

Genes Involved in Utilization of Methylated Amines

Three gene clusters, potentially involved in the utilization of TMA, DMA and MA as carbon, nitrogen and energy sources were identified in the *P. aminophilus* genome. Two of them are

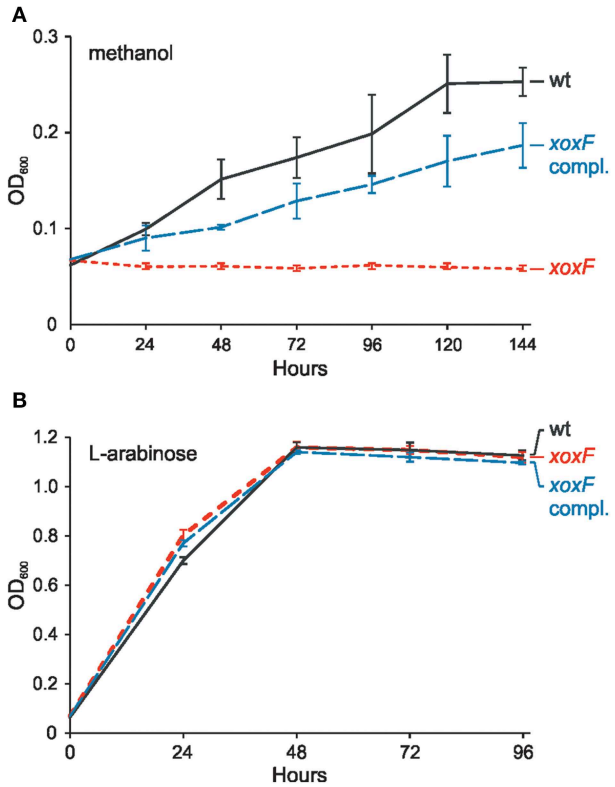


FIGURE 3 | Effect of the *xoxF* mutation on growth of *P. aminophilus* JCM 7686 on methanol (A) and L-arabinose (B). wt, wild type; *xoxF*, *xoxF* insertional mutant; *xoxF* compl., *xoxF* insertional mutant complemented with a copy of the *xoxF* gene cloned in pBBR1MCS-3. The values are means of three replicates, and the error bars indicate the standard deviations.

located within the chromid pAMI6 and contain genes encoding enzymes responsible for TMA oxidation via trimethylamine *N*-oxide (TMAO) and MA oxidation via *N*-methylglutamate. The third cluster is located within the chromosome and encodes methylamine dehydrogenase.

TMA oxidation via TMAO depends on the activity of three enzymes: (i) TMA monooxygenase, (ii) TMAO demethylase, and (iii) DMA monooxygenase. Interestingly, pAMI6 carries two genes, *tmm1* (JCM7686_pAMI6p076) and *tmm2* (JCM7686_pAMI6p102), encoding putative TMA monooxygenases. The predicted Tmm1 and Tmm2 proteins show high amino acid (aa) sequence similarity (75%) to one another and share about 60% identity with the Tmm protein of *Methylocella silvestris* BL2 (Chen et al., 2011). To analyze the role of the *tmm* genes in the TMA metabolism of JCM 7686, three mutant strains were constructed lacking either *tmm1* or *tmm2*, or both genes. The growth rate of the strains containing single mutations (*tmm1* or *tmm2*) in minimal medium with TMA as the sole carbon and energy source was identical to that of the wild type strain (**Figure 4**). In contrast, the double mutant strain (*tmm1tmm2*) was no longer able to utilize TMA, but it still had the ability to grow on medium supplemented with DMA or MA (the products of TMA utilization) (**Figure 4**). The growth on TMA was restored when *tmm1* or *tmm2* gene cloned in vector pBBR1MCS-5 was introduced into the double mutant strain (data not shown). These results indicated that both identified *tmm* genes encode enzymes with the same specificity and that they are both involved in the first stage of TMA metabolism. Additionally, the contribution of two *tmm* genes in TMA oxidation was confirmed by RT-qPCR. It was shown that transcript levels of both genes are increased in a similar degree during growth on TMA in comparison with non-methylotrophic conditions (**Table 1**).

Genes encoding TMAO demethylase (*tdm*, JCM7686_pAMI6p069) and DMA monooxygenase (*dmmDABC*, JCM7686_pAMI6p074-71) were identified in the vicinity of *tmm1*. The predicted Tdm of *P. aminophilus* shares 63% aa sequence identity with the Tdm protein of *Ruegeria pomeroyi* DSS-3 (Lidbury et al., 2014) and the DmmDABC proteins are homologous to four subunits of the putative DMA monooxygenase of *Methylocella silvestris* BL2 (Zhu et al., 2014). The DmmC proteins of strains JCM 7686 and BL2 share 66% aa sequence identity, while the three other putative DMA subunits encoded by these strains (DmmA, DmmB, and DmmD) are less well-conserved (39, 41, and 37% aa sequence identity, respectively).

To verify the function of the *P. aminophilus dmmDABC* genes, their mutational analysis was performed. The four mutant strains carrying deletions of the individual *dmm* genes failed to grow on dimethylamine as the sole carbon and energy source, while they showed the same growth rate as the wild-type strain when cultivated on methylamine or L-arabinose (**Figure 5**). The wild-type phenotype was restored when the *dmmDABC* module cloned in vector pBBR1MCS-3 was introduced into the mutant strains (data not shown). The mutations also influenced growth on C1 compounds that are metabolized via DMA, i.e., *N,N*-dimethylformamide and trimethylamine (**Figure 5**). Interestingly, inactivation of *dmmD* had a much weaker effect on growth on TMA than the inactivation of the other *dmm* genes (**Figure 5**). This observation is consistent with the hypothesis that the DmmD protein is not necessary for the conversion of DMA into MA and formaldehyde, but is an auxiliary subunit of the DMA monooxygenase which may convert formaldehyde into methylene-THF (Zhu et al., 2014).

Methylamine (MA), which is the product of DMA utilization, is presumably oxidized by *P. aminophilus* via two different pathways. Upstream and downstream of *tmm2*, we identified genes responsible for MA oxidation via *N*-methylglutamate (the *N*-methylglutamate pathway for MA oxidation, NMGP). They are arranged in two putative operons encoding three enzymes of the pathway: (i) the *mgsABC-gmaS* operon (JCM7686_pAMI6p108-105) encoding glutamate-MA ligase (GmaS) and *N*-methyl-L-glutamate synthase (MgsABC), and (ii) the *mgdABCD* operon (JCM7686_pAMI6p099-096) encoding *N*-methyl-L-glutamate dehydrogenase (MgdABCD). Most of the NMGP enzymes subunits show high levels of aa sequence identity with the respective proteins of *M. silvestris* BL2 (Chen et al., 2010).

The second *P. aminophilus* pathway responsible for MA oxidation relies on the activity of a two-subunit MA dehydrogenase (Mau). Genes encoding this enzyme are

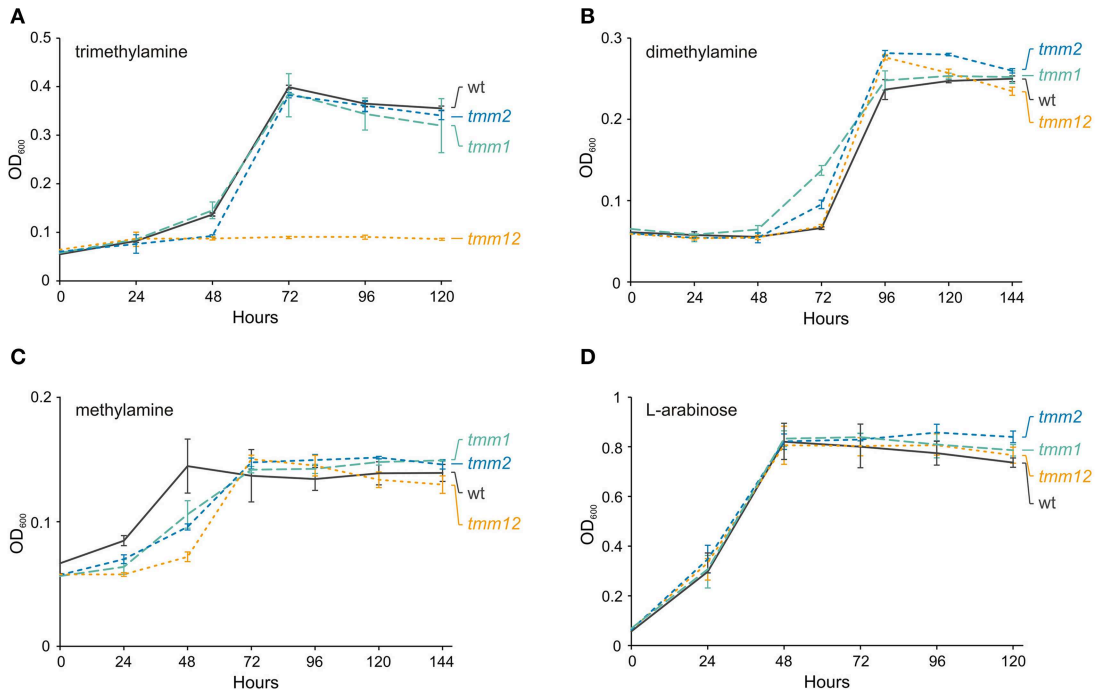


FIGURE 4 | Effect of the *tmm1* and *tmm2* mutations on growth of *P. aminophilus* JCM 7686 on trimethylamine (A), dimethylamine (B), methylamine (C), and L-arabinose (D). wt, wild type; *tmm1*, *tmm1* insertional mutant; *tmm2*, *tmm2* insertional mutant; *tmm12*, double insertional mutant in *tmm1* and *tmm2* genes. The values are means of three replicates, and the error bars indicate the standard deviations.

TABLE 1 | Changes in the transcript levels of selected methylotrophy genes of *P. aminophilus* JCM 7686 in methylotrophic vs. non-methylotrophic conditions determined by RT-qPCR analysis.

ORF name	Protein name	Process	Fold change ± SD*
tmm1	TMA monooxygenase	oxidation of TMA	26.7 ± 2.8
tmm2	TMA monooxygenase	oxidation of TMA	27.3 ± 1.9
hpr	hydroxypyruvate reductase	serine cycle	3.0 ± 1.0
gck	glycerate 2-kinase	serine cycle	5.9 ± 0.7
aceA	isocitrate lyase	glyoxylate shunt	13.5 ± 2.6
ecm	ethylmalonyl-CoA mutase	ethylmalonyl-CoA pathway	4.9 ± 0.6
mcm	methylmalonyl-CoA mutase	ethylmalonyl-CoA pathway	4.6 ± 1.0

*SD, standard deviation.

arranged in a large chromosomal cluster (JCM7686_0162-0171, *mauFBEDACJGMN*). This gene cluster is very similar to the *mau* region located in plasmid 1 of *P. denitrificans* Pd1222 (approximately 75% nucleotide sequence identity) (van Der Palen et al., 1995). The *mau* genes are crucial for growth of *P. denitrificans* Pd1222 on MA (van Der Palen et al., 1995). However, in *P. aminophilus* we found that mutation of the gene encoding the catalytic subunit MauA did not abolish growth on this compound. In this case, the growth rate on MA-containing medium was reduced (**Figure S2**), which indicated that both identified pathways for MA utilization are active in *P. aminophilus*.

Genes Involved in Utilization of Formamides

Strain JCM 7686 can also utilize formamides, including *N,N*-dimethylformamide (**Figures 1**, **2**). Formamide breakdown is most probably catalyzed by two chromosomally- and pAMI4-encoded acetamidases/formamidases (JCM7686_1450, JCM7686_pAMI4p036), while DMF utilization is dependent on a pAMI2-encoded *N,N*-dimethylformamidase (DMFase) (JCM7686_pAMI2p015-017). The latter assumption was confirmed in our previous study in which we found that *P. aminophilus* strain deprived of pAMI2 was unable to utilize DMF (Dziewit et al., 2010). Plasmid pAMI2 carries genes encoding two subunits of DMFase (DmfA1 and DmfA2) which are organized in an operon. More detailed studies revealed that the expression of these genes is activated in the presence of DMF by the LuxR-family transcriptional activator DmfR (Dziewit et al., 2010). Interestingly, a related *dmfA1-dmfA2* locus (encoding predicted proteins sharing 34% and 40% aa sequence identity with DmfA1 and DmfA2 of pAMI2, respectively) was also identified within chromid pAMI5 (JCM7686_pAMI5p063-064). However, these genes have not been found to be linked to DMF utilization.

Genes Involved in Oxidation of Formaldehyde and CH_2 =THF to CO_2 and Reduction of Formate to CH_2 =THF

Most of the aforementioned pathways of C1 substrate oxidation result in the formation of formaldehyde and/or CH_2=THF (Chistoserdova, 2011; Keltjens et al., 2014; Lidbury et al.,

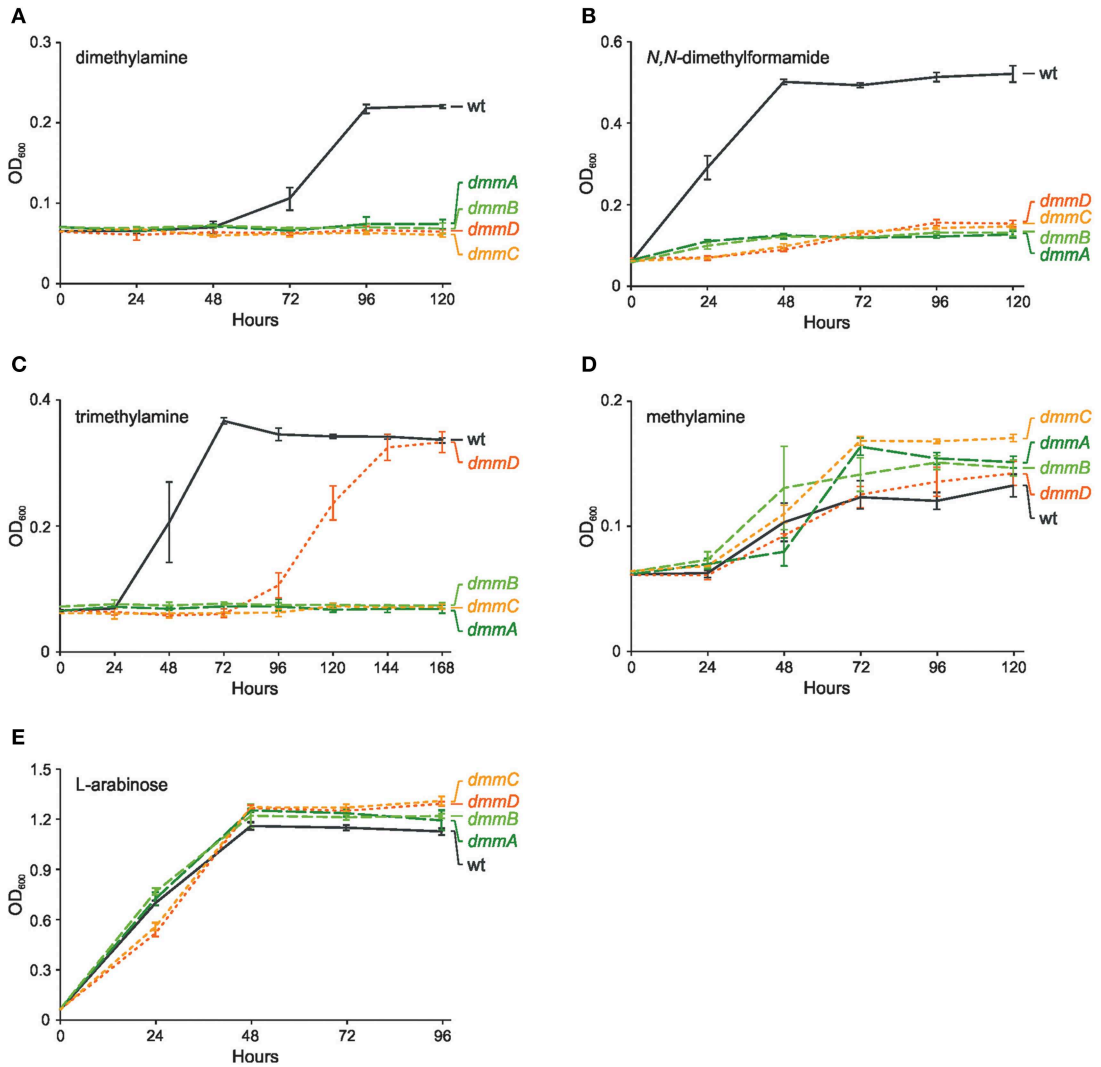


FIGURE 5 | Effect of mutations in genes encoding four putative subunits of dimethylamine monooxygenase (*dmmDABC*) on growth of *P. aminophilus* JCM 7686 on dimethylamine (A), *N,N*-dimethylformamide (B), trimethylamine (C), methylamine (D), and L-arabinose (E). wt, wild type; *dmmA*, *dmmA* insertional mutant; *dmmB*, *dmmB* insertional mutant; *dmmC*, *dmmC* insertional mutant; *dmmD*, *dmmD* insertional mutant. The values are means of three replicates, and the error bars indicate the standard deviations.

2014; Nayak and Marx, 2014) (**Figure 2**). In *P. aminophilus* JCM 7686, the oxidation of formaldehyde to formate is most likely achieved via the glutathione-dependent pathway utilizing three chromosomally encoded enzymes: (i) *S*-(hydroxymethyl)glutathione synthase (Gfa, JCM7686_0085), (ii) *S*-(hydroxymethyl)glutathione dehydrogenase (FlhA, JCM7686_0086), and (iii) *S*-formylglutathione hydrolase (FghA, JCM7686_0089). Each of these enzymes shares a high level of aa sequence identity (~85%) with the corresponding protein from *P. denitrificans* Pd1222. It is worth noting that the *S*-(hydroxymethyl)glutathione dehydrogenase was previously recognized as essential for methylotrophic growth of *P. denitrificans* (Ras et al., 1995).

P. aminophilus JCM 7686 encodes enzymes involved in transitions between CH_2=THF and formate in both the oxidizing and reducing directions. The oxidation of the methylene group of CH_2=THF is associated with energy release, while the reductive pathway is required to supply CH_2=THF to the serine cycle (see below) during growth of *P. aminophilus* on methanol and formamide (i.e., C1 compounds, whose oxidation leads to the formation of formaldehyde or formate but not CH_2=THF) (**Figure 2**).

Oxidation of the methylene group to formate is performed by the action of two enzymes: CH_2=THF dehydrogenase/CH_2=THF cyclohydrolase (FolD) and formyltetrahydrofolate deformylase (PurU). FolD and PurU are encoded by three homologous two-gene loci (*folD-purU*) located in pAMI6 (2 loci) and pAMI5 (**Table S4**). In the reduction pathway, formate is loaded into THF by formate-tetrahydrofolate ligase (FtfL, JCM7686_pAMI6p042) and then

reduced to the methylene group. In many methylotrophs (e.g., *Methylobacterium extorquens*) the reduction process is performed by the sequential action of two enzymes: CH_2=THF cyclohydrolase (Fch) and CH_2=THF dehydrogenase (MtdA) (Chistoserdova, 2011). Since *P. aminophilus* encodes neither Fch nor MtdA, the reduction of 10-formyl-THF to CH_2=THF is presumably catalyzed by FolD which seems to work bi-directionally in some bacteria (Beck et al., 2015).

The product of oxidation of formaldehyde and the methylene group of CH_2=THF is formate. It is further oxidized to CO_2 by formate dehydrogenase (Fdh) (Chistoserdova et al., 2004). In the *P. aminophilus* genome there are four gene clusters encoding formate dehydrogenases; three of them are located in the chromosome (JCM7686_0639-0643, JCM7686_2088, JCM7686_3476-3480) and one in plasmid pAMI1 (JCM7686_pAMI1p027).

Genes Involved in Assimilation of C1 Units

In contrast to other methylotrophic strains of the genus *Paracoccus*, *P. aminophilus* JCM 7686 is unable to grow autotrophically using the Calvin cycle. The genome of this strain does not encode subunits of the key enzyme of this process, ribulose-1,5-bisphosphate carboxylase/oxygenase (RuBisCO). Therefore, the only possible way to assimilate C1 compounds seems to be via the serine cycle. *P. aminophilus* carries a chromosomally encoded cluster of serine cycle genes, which is also highly conserved in the genome of *Paracoccus* sp. N5 (Beck et al., 2015). Since many of these genes encode proteins that are highly divergent from the well-studied serine cycle enzymes of other methylotrophs (Beck et al., 2015), functional analysis was required to confirm their specific activities.

The serine cycle gene cluster of *P. aminophilus* is incomplete since it does not contain a *gck* gene encoding glycerate 2-kinase, which is, surprisingly, present at a different genomic location—within chromid pAMI5. Interestingly, the Gck of *P. aminophilus* is more closely related to a protein from the marine bacterium *Labrenzia alexandrii* DFL-11 (72% aa sequence identity) than to the Gck of *Paracoccus* sp. N5 (38% aa sequence identity) and other homologous genes found in *Paracoccus* spp., which strongly suggests that the JCM 7686 *gck* gene was independently acquired by horizontal gene transfer.

The serine cycle cannot operate without regeneration of glyoxylate from acetyl-CoA, which can proceed via the glyoxylate shunt or the ethylmalonyl-CoA pathway (EMCP) (Chistoserdova, 2011). Both pathways also enable growth on C2 compounds. The chromosome of *P. aminophilus* contains the genetic information required for the synthesis of all enzymes of the EMCP (**Table S4**). Moreover, pAMI6 contains a two-gene locus (JCM7686_pAMI6p120-121) encoding putative enzymes of the glyoxylate shunt: isocitrate lyase (AceA) and malate synthase (AceB). An additional copy of the gene encoding malate synthase (malate synthase G, GlcB, JCM7686_1627) was identified within the chromosome of this strain. Therefore, our bioinformatic sequence analysis indicates that *P. aminophilus* may regenerate glyoxylate via both the EMCP and glyoxylate shunt.

We performed RT-qPCR analyses to confirm that the predicted serine cycle genes (*hpr* carried within pAMI6 and *gck* carried within pAMI5), the gene of isocitrate lyase (*aceA*) involved in the glyoxylate shunt, as well as two genes of the ethylmalonyl-CoA pathway (*ecm* and *mcm*) are linked to methylotrophic metabolism of the JCM 7686 strain. The results showed that in each case the transcript level was elevated during methylotrophic growth (**Table 1**).

Abundance of Methylotrophy-linked Genes in *Paracoccus* spp. Genomes

Annotated genomic sequences of 17 strains of *Paracoccus* spp. (including *P. aminophilus* JCM 7686) are currently available in the GenBank database (**Figure 6**). The genomes of these strains were screened for the presence of 67 genes encoding enzymes involved in methylotrophy that have been identified in *P. aminophilus* and other methylotrophic *Alphaproteobacteria*. It is important to note that for 15 *Paracoccus* spp. strains only draft genomes were available, thus some data may be missing. Therefore, the lack of the particular genes has to be verified after obtaining complete genome sequences of those strains.

On the basis of the performed analysis, strains of *Paracoccus* spp. were classified into four groups: (i) autotrophic methylotrophs, which can potentially assimilate CO_2 via the Calvin cycle (six strains), (ii) facultatively autotrophic methylotrophs, which have both the serine cycle and the Calvin cycle (3 strains), (iii) heterotrophic methylotrophs, which use only the serine pathway for carbon assimilation, represented only by *P. aminophilus* JCM 7686, and (iv) non-methylotrophs (**Figure 6**).

More differences were observed when the strains were compared in terms of the range of C1 compounds that potentially can be oxidized. All but one strain (*P. halophilus* JCM 14014) encode XoxF methanol dehydrogenase, and six strains (*P. denitrificans* Pd1222, *P. pantotrophus* J40, J46 and *Paracoccus* spp. J39, J55, TRP) also encode a PQQ-dependent calcium-binding methanol dehydrogenase (MxaFI) (**Figure 6**). Eight of the analyzed strains encode enzymes responsible for the utilization of methylated amines. Of these, five (JCM 7686, J39, J55, N5, and ATCC BAA-599) also encode enzymes that are required for trimethylamine oxidation via trimethylamine *N*-oxide. In addition, these five strains plus *P. versutus* DSM 582 possess *dmmABCD* genes encoding putative dimethylamine monooxygenase (**Figure 6**).

Another interesting observation was made while analyzing genes involved in methylamine utilization. Three groups of strains were distinguished that are likely to be able to perform methylamine oxidation via different routes: (i) the NMGP (all facultatively autotrophic methylotrophs, i.e., *Paracoccus* sp. J39, J55, and N5), (ii) the pathway involving Mau (*P. denitrificans* Pd1222 and *Paracoccus* sp. TRP), and (iii) both pathways (*P. aminophilus*, *P. yeei,* and *P. versutus*) (**Figure 6**).

What is noteworthy, in all strains possessing the serine cycle genes, a co-occurrence of the genes involved in TMA oxidation via TMAO and the NMGP was observed. In *P. aminophilus* JCM 7686, these genes are present within three separate clusters. Two of them are located in pAMI6 (genes involved in TMA oxidation and in the NMGP) and the other (involved in the serine cycle) in the chromosome (**Figure 7**). In *Paracoccus* sp.

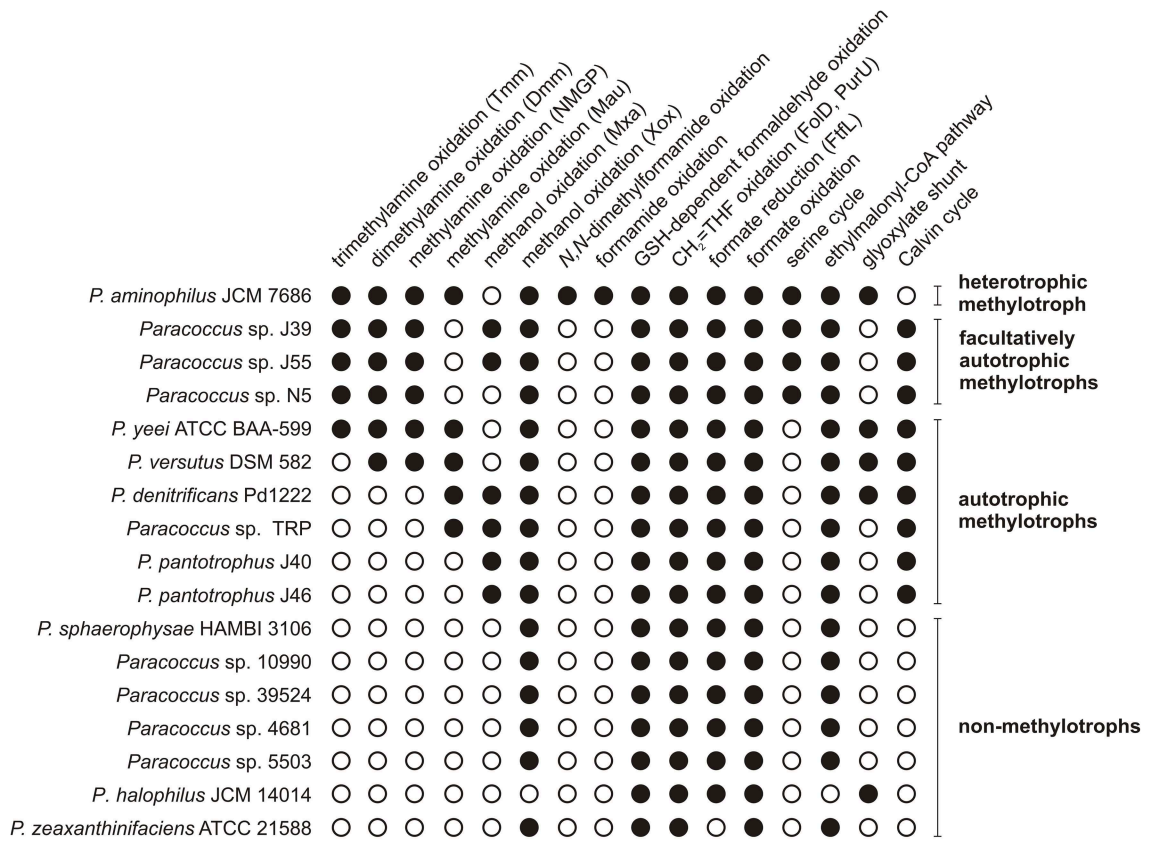


FIGURE 6 | The methylotrophy-linked metabolic capabilities encoded by the genomes of *Paracoccus* spp. Tmm, trimethylamine monooxygenase; Dmm, putative dimethylamine monooxygenase; NMGP, *N*-methylglutamate pathway; Mau, methylamine dehydrogenase; Mxa, Mxa-type methanol dehydrogenase; Xox, Xox-type methanol dehydrogenase; GSH, glutathione; CH_2=THF, 5,10-methylene-tetrahydrofolate; FolD, 5,10-methylene-tetrahydrofolate dehydrogenase/5,10-methylene-tetrahydrofolate cyclohydrolase; PurU, formyltetrahydrofolate deformylase; FtfL, formate-tetrahydrofolate ligase. The accession numbers of *Paracoccus* complete genomes: JCM7686 – GQ410978, GQ468939, GQ468938, CP006650–CP006655, and Pd1222 – CP000489–CP000491. The accession numbers of *Paracoccus* draft genomes: J39 – JAEN01000001–JAEN01000050, J55 – AZVA01000001–AZVA01000069, N5 – AQUO01000001–AQUO01000003, ATCC BAA-599 – JHWH01000001– JHWH01000073, DSM 582 – JRKO01000001– JRKO01000187, TRP – AEPN01000001– AEPN01000119, J40 – JAGK01000001– JAGK01000119, J46 – JAEM01000001– JAEM01000105, HAMBI 3106 – JRKS01000001– JRKS01000137, 10990 – JRKR01000001– JRKR01000309, 39524 – JRKP01000001– JRKP01000273, 4681 – JRKT01000001– JRKT01000176, 5503 – JRKQ01000001– JRKQ01000265, JCM 14014 – JRKN01000001– JRKN01000122, and ATCC 21588 – ATUJ01000001– ATUJ01000035.

N5 all the aforementioned genes form a single methylotrophy island, which is most probably extrachromosomally located (**Figure 7**).

Interestingly, we found that the gene clusters homologous with the methylotrophy island of *Paracoccus* sp. N5 are also present in the chromosomes of numerous strains of the marine *Roseobacter* clade (**Figure 7**), many of which were recognized as serine cycle methylotrophs (Newton et al., 2010). These bacteria comprise up to 20% of the microorganisms in coastal surface waters (Chen, 2012). Therefore, the identified methylotrophy islands may constitute one of the most abundant sets of genes participating in C1 metabolism worldwide. Since these islands contain genes involved in all three stages of the methylotrophy process (oxidation of specific C1 substrates, oxidation of CH_2=THF and assimilation of C1 units), their transfer to other hosts may result in the conversion of non-methylotrophic strains into methylotrophs. It is noteworthy that bacteria of the *Roseobacter* clade are well known for the presence of numerous conjugative megaplasmids, which may promote horizontal transmission of large segments of genomic DNA (Petersen et al., 2012, 2013).

In the case of *P. aminophilus* JCM 7686 several other extrachromosomal elements besides pAMI6 (plasmids pAMI1, pAMI2, pAMI4, and chromid pAMI5) contain genes linked to methylotrophy. As previously shown (Dziewit et al., 2014), the host range of these elements is not limited to *Paracoccus* spp. and extends to other strains of *Alphaproteobacteria*, including *Agrobacterium tumefaciens* and *Rhizobium etli*. Thus, further transfer of these extrachromosomally-located methylotrophy genes to other hosts may result in the formation of "patchwork" methylotrophic pathways and the generation of bacterial strains with novel metabolic properties.

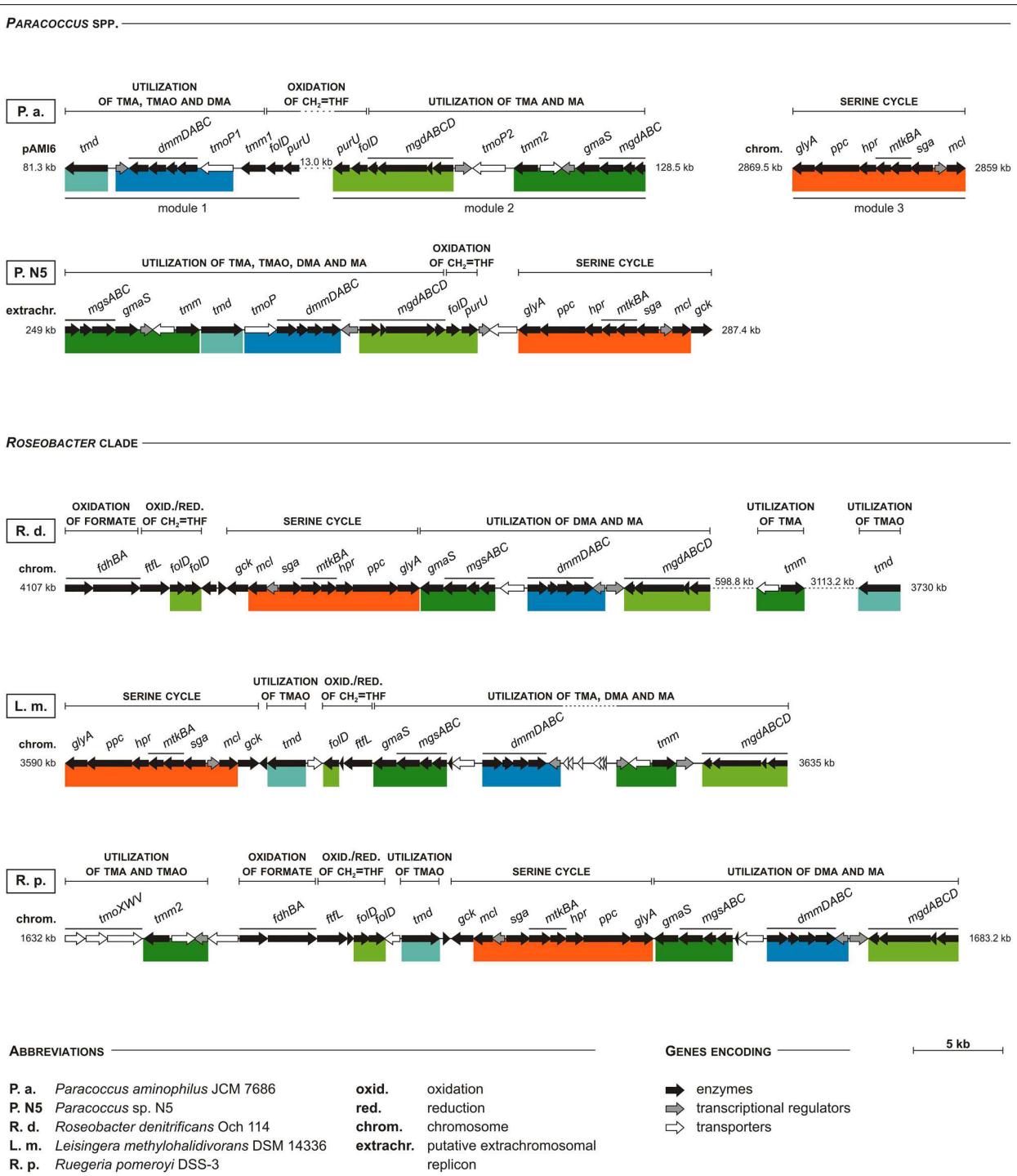


FIGURE 7 | Comparison of clustering of the serine pathway genes with other methylotrophy-linked genes in *P. aminophilus* JCM 7686, *Paracoccus* sp. N5, and three selected strains of the *Roseobacter* clade. Colored bars show shuffling of specific DNA regions when compared to *P. aminophilus* JCM 7686. In *P. aminophilus* JCM 7686, the presented set of methylotrophy genes is divided into three clusters, two located within pAMI6 and one within the chromosome. In the other strains all these genes constitute a single methylotrophy island, which may be located within the chromosome (*R. pomeroyi* DSS-3 and *L. methylohalidivorans* DSM 14336), or in putative extrachromosomal replicons (*Paracoccus* sp. N5). In some cases, certain genes of the methylotrophy island may be found in a distant location in a genome (e.g., the *tmm* and *tdm* genes of *R. denitrificans* Och 114). The full names of enzymes encoded by most of the genes presented in the figure are summarized in **Table S4**. The others are *tmoP*—TMAO permease gene (Zhu et al., 2014), and *tmoXWV*—TMAO ABC-transporter system genes (Lidbury et al., 2014). The accession numbers of particular replicons are as follows: JCM 7686 – CP006654 (pAMI6) and CP006650 (chromosome), N5 – AQUO01000003, Och 114 – CP000362, DSM 14336 – CP006773, DSS-3 – CP000031.

Conclusions

The major goal of this study was to define the metabolic network involved in the C1 metabolism of *P. aminophilus* JCM 7686 and to compare it at the genetic level with other members of the genus *Paracoccus*. A genome-wide analysis revealed the great methylotrophic potential of this strain, manifested in its ability to utilize a wide range of C1 compounds, including formamide and *N*,*N*-dimethylformamide. These phenotypes enabled adaptation of *P. aminophilus* to its natural "methylotrophic" habitat, which was soil contaminated with DMF.

On the basis of the *in silico* and experimental evidence we defined complex methylotrophic pathways of *P. aminophilus*, with by-pass pathways for methylamine utilization and genes encoding multiple enzymes of the same specificity involved e.g., in trimethylamine utilization. The analysis revealed that this strain is the only known heterotrophic methylotroph among *Paracoccus* spp., encoding enzymes of the serine cycle as an exclusive pathway for C1 unit assimilation. Thus, *P. aminophilus* JCM 7686 is an excellent model for the studies on genetic diversity and evolution of methylotrophy in this group of bacteria.

Acknowledgments

We acknowledge J. Baj for helpful discussions and advice, L. Kowalski and K. Kuzmicz for technical assistance, and M. Puzyna for help in the preparation of the figures. This work was founded by the Ministry of Science and Higher Education, Poland (grant: IP2011 011471) and National Science Centre (NCN), Poland, on the basis of the decision number DEC-2013/09/B/NZ1/00133. The bioinformatics support of the BMBF-funded project "Bielefeld-Gießen Center for Microbial Bioinformatics—BiGi (grant: 031A533)" within the German Network for Bioinformatics Infrastructure (de.NBI) is gratefully acknowledged.

Supplementary Material

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fmicb.2015.00852

Figure S1 | Chemical structures of C1 compounds utilized by *P. aminophilus* JCM 7686.

Figure S2 | Effect of the *mauA* mutation on growth of *P. aminophilus* JCM 7686 on methylamine. wt, wild type; *mauA*, *mauA* insertional mutant. The values are means of three replicates, and the error bars indicate the standard deviations.

Table S1 | Bacterial strains used in this study.

Table S2 | Plasmids used and constructed in this study.

Table S3 | Oligonucleotide primers used in this study.

Table S4 | Genes of *P. aminophilus* JMC 7686 involved in the metabolism of C1 compounds.

References

Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., et al. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25, 3389–3402. doi: 10.1093/nar/25.17.3389

Baker, S. C., Ferguson, S. J., Ludwig, B., Page, M. D., Richter, O. M., and van Spanning, R. J. (1998). Molecular genetics of the genus *Paracoccus*: metabolically versatile bacteria with bioenergetic flexibility. *Microbiol. Mol. Biol. Rev.* 62, 1046–1078.

Bartosik, D., Baj, J., Bartosik, A. A., and Wlodarczyk, M. (2002). Characterization of the replicator region of megaplasmid pTAV3 of *Paracoccus versutus* and search for plasmid-encoded traits. *Microbiology* 148, 871–881. doi: 10.1099/00221287-148-3-871

Bartosik, D., Szymanik, M., and Wysocka, E. (2001). Identification of the partitioning site within the *repABC*-type replicon of the composite *Paracoccus versutus* plasmid pTAV1. *J. Bacteriol.* 183, 6234–6243. doi: 10.1128/JB.183.21.6234-6243.2001

Beck, D. A., McTaggart, T. L., Setboonsarng, U., Vorobev, A., Goodwin, L., Shapiro, N., et al. (2015). Multiphyletic origins of methylotrophy in *Alphaproteobacteria*, exemplified by comparative genomics of Lake Washington isolates. *Environ. Microbiol.* 17, 547–554. doi: 10.1111/1462-2920.12736

Caspi, R., Foerster, H., Fulcher, C. A., Kaipa, P., Krummenacker, M., Latendresse, M., et al. (2008). The MetaCyc Database of metabolic pathways and enzymes and the BioCyc collection of pathway/genome databases. *Nucleic Acids Res.* 36, D623–D631. doi: 10.1093/nar/gkm900

Chen, Y. (2012). Comparative genomics of methylated amine utilization by marine *Roseobacter* clade bacteria and development of functional gene markers (*tmm*, *gmaS*). *Environ. Microbiol.* 14, 2308–2322. doi: 10.1111/j.1462-2920.2012.02765.x

Chen, Y., Patel, N. A., Crombie, A., Scrivens, J. H., and Murrell, J. C. (2011). Bacterial flavin-containing monooxygenase is trimethylamine monooxygenase. *Proc. Natl. Acad. Sci. U.S.A.* 108, 17791–17796. doi: 10.1073/pnas.1112928108

Chen, Y., Scanlan, J., Song, L., Crombie, A., Rahman, M. T., Schäfer, H., et al. (2010). γ-Glutamylmethylamide is an essential intermediate in the metabolism of methylamine by *Methylocella silvestris*. *Appl. Environ. Microbiol.* 76, 4530–4537. doi: 10.1128/AEM.00739-10

Chistoserdova, L. (2011). Modularity of methylotrophy, revisited. *Environ. Microbiol.* 13, 2603–2622. doi: 10.1111/j.1462-2920.2011.02464.x

Chistoserdova, L., Laukel, M., Portais, J. C., Vorholt, J. A., and Lidstrom, M. E. (2004). Multiple formate dehydrogenase enzymes in the facultative methylotroph *Methylobacterium extorquens* AM1 are dispensable for growth on methanol. *J. Bacteriol.* 186, 22–28. doi: 10.1128/JB.186.1.22-28.2004

Claudel-Renard, C., Chevalet, C., Faraut, T., and Kahn, D. (2003). Enzyme-specific profiles for genome annotation: PRIAM. *Nucleic Acids Res.* 31, 6633–6639. doi: 10.1093/nar/gkg847

Ditta, G., Stanfield, S., Corbin, D., and Helinski, D. R. (1980). Broad host range DNA cloning system for gram-negative bacteria: construction of a gene bank of *Rhizobium meliloti*. *Proc. Natl. Acad. Sci. U.S.A.* 77, 7347–7351. doi: 10.1073/pnas.77.12.7347

Dziewit, L., Adamczuk, M., Szuplewska, M., and Bartosik, D. (2011a). DIY series of genetic cassettes useful in construction of versatile vectors specific for *Alphaproteobacteria*. *J. Microbiol. Methods* 86, 166–174. doi: 10.1016/j.mimet.2011.04.016

Dziewit, L., Czarnecki, J., Wibberg, D., Radlinska, M., Mrozek, P., Szymczak, M., et al. (2014). Architecture and functions of a multipartite genome of the methylotrophic bacterium *Paracoccus aminophilus* JCM 7686, containing primary and secondary chromids. *BMC Genomics* 15:124. doi: 10.1186/1471-2164-15-124

Dziewit, L., Dmowski, M., Baj, J., and Bartosik, D. (2010). Plasmid pAMI2 of *Paracoccus aminophilus* JCM 7686 carries *N,N*-dimethylformamide degradation-related genes whose expression is activated by a LuxR family regulator. *Appl. Environ. Microbiol.* 76, 1861–1869. doi: 10.1128/AEM.01926-09

Dziewit, L., Jazurek, M., Drewniak, L., Baj, J., and Bartosik, D. (2007). The SXT conjugative element and linear prophage N15 encode toxin-antitoxin-stabilizing systems homologous to the *tad-ata* module of the *Paracoccus aminophilus* plasmid pAMI2. *J. Bacteriol.* 189, 1983–1997. doi: 10.1128/JB.01610-06

Dziewit, L., Kuczkowska, K., Adamczuk, M., Radlinska, M., and Bartosik, D. (2011b). Functional characterization of the type II PamI restriction-modification system derived from plasmid pAMI7 of *Paracoccus aminophilus* JCM 7686. *FEMS Microbiol. Lett.* 324, 56–63. doi: 10.1111/j.1574-6968.2011.02388.x

Funke, G., Frodl, R., and Sommer, H. (2004). First comprehensively documented case of *Paracoccus yeei* infection in a human. *J. Clin. Microbiol.* 42, 3366–3368. doi: 10.1128/JCM.42.7.3366-3368.2004

Gibson, T. J. (1984). *Studies on Epstein-Barr Genome.* Ph.D. thesis, University of Cambridge.

Heinl, S., Wibberg, D., Eikmeyer, F., Szczepanowski, R., Blom, J., Linke, B., et al. (2012). Insights into the completely annotated genome of *Lactobacillus buchneri* CD034, a strain isolated from stable grass silage. *J. Biotechnol.* 161, 153–166. doi: 10.1016/j.jbiotec.2012.03.007

Karp, P. D., Paley, S. M., Krummenacker, M., Latendresse, M., Dale, J. M., Lee, T. J., et al. (2010). Pathway Tools version 13.0: integrated software for pathway/genome informatics and systems biology. *Brief. Bioinform.* 11, 40–79. doi: 10.1093/bib/bbp043

Kelly, D. P., Rainey, F. A., and Wood, A. P. (2006). "The genus Paracoccus," in *Prokaryotes,* eds M. Dworkin, S. Falkow, E. Rosenberg, K.-H. Schleifer and E. Stackebrandt (New York, NY: Springer), 232–249.

Keltjens, J. T., Pol, A., Reimann, J., and Op den Camp, H. J. (2014). PQQ-dependent methanol dehydrogenases: rare-earth elements make a difference. *Appl. Microbiol. Biotechnol.* 98, 6163–6183. doi: 10.1007/s00253-014-5766-8

Kim, S. G., Bae, H. S., and Lee, S. T. (2001). A novel denitrifying bacterial isolate that degrades trimethylamine both aerobically and anaerobically via two different pathways. *Arch. Microbiol.* 176, 271–277. doi: 10.1007/s002030100319

Kim, S. G., Bae, H. S., Oh, H. M., and Lee, S. T. (2003). Isolation and characterization of novel halotolerant and/or halophilic denitrifying bacteria with versatile metabolic pathways for the degradation of trimethylamine. *FEMS Microbiol. Lett.* 225, 263–269. doi: 10.1016/S0378-1097(03)00530-5

Kovach, M. E., Phillips, R. W., Elzer, P. H., Roop, R. M. II., and Peterson, K. M. (1994). pBBR1MCS: a broad-host-range cloning vector. *BioTechniques* 16, 800–802.

Kushner, S. R. (1978). "An improved method for transformation of *E. coli* with ColE1 derived plasmids," in *Genetic Engineering,* eds H. B. Boyer and S. Nicosia (Amsterdam: Elsevier/North-Holland), 17–23.

Lee, J. H., Kim, Y. S., Choi, T. J., Lee, W. J., and Kim, Y. T. (2004). *Paracoccus haeundaensis* sp. nov., a Gram-negative, halophilic, astaxanthin-producing bacterium. *Int. J. Syst. Evol. Microbiol.* 54, 1699–1702. doi: 10.1099/ijs.0.63146-0

Li, K., Wang, S., Shi, Y., Qu, J., Zhai, Y., Xu, L., et al. (2011). Genome sequence of *Paracoccus* sp. Strain TRP, a chlorpyrifos biodegrader. *J. Bacteriol.* 193, 1786–1787. doi: 10.1128/JB.00014-11

Lidbury, I., Murrell, J. C., and Chen, Y. (2014). Trimethylamine N-oxide metabolism by abundant marine heterotrophic bacteria. *Proc. Natl. Acad. Sci. U.S.A.* 111, 2710–2715. doi: 10.1073/pnas.1317834111

Lipski, A., Reichert, K., Reuter, B., Sproer, C., and Altendorf, K. (1998). Identification of bacterial isolates from biofilters as *Paracoccus alkenifer* sp. nov. and *Paracoccus solventivorans* with emended description of *Paracoccus solventivorans. Int. J. Syst. Bacteriol.* 48, 529–536. doi: 10.1099/00207713-48-2-529

Liu, X. Y., Wang, B. J., Jiang, C. Y., and Liu, S. J. (2006). *Paracoccus sulfuroxidans* sp. nov., a sulfur oxidizer from activated sludge. *Int. J. Syst. Evol. Microbiol.* 56, 2693–2695. doi: 10.1099/ijs.0.64548-0

Liu, Y., Gan, L., Chen, Z., Megharaj, M., and Naidu, R. (2012). Removal of nitrate using *Paracoccus* sp. YF1 immobilized on bamboo carbon. *J. Hazard. Mater.* 229–230, 419–425. doi: 10.1016/j.jhazmat.2012.06.029

Meyer, F., Goesmann, A., McHardy, A. C., Bartels, D., Bekel, T., Clausen, J., et al. (2003). GenDB - an open source genome annotation system for prokaryote genomes. *Nucleic Acids Res.* 31, 2187–2195. doi: 10.1093/nar/gkg312

Nayak, D. D., and Marx, C. J. (2014). Methylamine utilization via the *N*-methylglutamate pathway in *Methylobacterium extorquens* PA1 involves a novel flow of carbon through C1 assimilation and dissimilation pathways. *J. Bacteriol.* 196, 4130–4139. doi: 10.1128/JB.02026-14

Newton, R. J., Griffin, L. E., Bowles, K. M., Meile, C., Gifford, S., Givens, C. E., et al. (2010). Genome characteristics of a generalist marine bacterial lineage. *ISME J.* 4, 784–798. doi: 10.1038/ismej.2009.150

Petersen, J., Brinkmann, H., Bunk, B., Michael, V., Päuker, O., and Pradella, S. (2012). Think pink: photosynthesis, plasmids and the *Roseobacter* clade. *Environ. Microbiol.* 14, 2661–2672. doi: 10.1111/j.1462-2920.2012.02806.x

Petersen, J., Frank, O., Göker, M., and Pradella, S. (2013). Extrachromosomal, extraordinary and essential–the plasmids of the *Roseobacter* clade. *Appl. Microbiol. Biotechnol.* 97, 2805–2815. doi: 10.1007/s00253-013-4746-8

Philippe, N., Alcaraz, J. P., Coursange, E., Geiselmann, J., and Schneider, D. (2004). Improvement of pCVD442, a suicide plasmid for gene allele exchange in bacteria. *Plasmid* 51, 246–255. doi: 10.1016/j.plasmid.2004.02.003

Ras, J., Van Ophem, P. W., Reijnders, W. N., Van Spanning, R. J., Duine, J. A., Stouthamer, A. H., et al. (1995). Isolation, sequencing, and mutagenesis of the gene encoding NAD- and glutathione-dependent formaldehyde dehydrogenase (GD-FALDH) from *Paracoccus denitrificans*, in which GD-FALDH is essential for methylotrophic growth. *J. Bacteriol.* 177, 247–251.

Reece, K. S., and Phillips, G. J. (1995). New plasmids carrying antibiotic-resistance cassettes. *Gene* 165, 141–142. doi: 10.1016/0378-1119(95)00529-F

Sambrook, J., and Russell, D. W. (2001). *Molecular Cloning: A Laboratory Manual.* New York, NY: Cold Spring Harbor Laboratory Press.

Sanjeevkumar, S., Nayak, A. S., Santoshkumar, M., Siddavattam, D., and Karegoudar, T. B. (2013). *Paracoccus denitrificans* SD1 mediated augmentation with indigenous mixed cultures for enhanced removal of *N,N*-dimethylformamide from industrial effluents. *Biochem. Eng. J.* 79, 1–6. doi: 10.1016/j.bej.2013.06.016

Siller, H., Rainey, F. A., Stackebrandt, E., and Winter, J. (1996). Isolation and characterization of a new gram-negative, acetone-degrading, nitrate-reducing bacterium from soil, *Paracoccus solventivorans* sp. nov. *Int. J. Syst. Bacteriol.* 46, 1125–1130. doi: 10.1099/00207713-46-4-1125

Simon, R., Priefer, U., and Pühler, A. (1983). A broad host range mobilization system for *in vivo* genetic-engineering - transposon mutagenesis in gram-negative bacteria. *Bio-Technology* 1, 784–791. doi: 10.1038/nbt1183-784

Sun, M., Luo, Y., Teng, Y., Christie, P., Jia, Z., and Li, Z. (2013). Tenax TA extraction to understand the rate-limiting factors in methyl-beta-cyclodextrin-enhanced bioremediation of PAH-contaminated soil. *Biodegradation* 24, 365–375. doi: 10.1007/s10532-012-9593-2

Trotsenko, Y. A., and Murrell, J. C. (2008). Metabolic aspects of aerobic obligate methanotrophy. *Adv. Appl. Microbiol.* 63, 183–229. doi: 10.1016/S0065-2164(07)00005-6

Tsubokura, A., Yoneda, H., and Mizuta, H. (1999). *Paracoccus carotinifaciens* sp. nov., a new aerobic gram-negative astaxanthin-producing bacterium. *Int. J. Syst. Bacteriol.* 49(Pt 1), 277–282. doi: 10.1099/00207713-49-1-277

Turova, T. P., Kuznetsov, B. B., Doronina, N. V., and Trotsenko Iu, A. (2001). Phylogenetic analysis of aerobic methylotrophic bacteria, using dichloromethane. *Mikrobiologiia* 70, 92–97. doi: 10.1023/A:1004801106488

Urakami, T., Araki, H., Oyanagi, H., Suzuki, K., and Komagata, K. (1990). *Paracoccus aminophilus* sp. nov. and *Paracoccus aminovorans* sp. nov., which utilize N,N-dimethylformamide. *Int. J. Syst. Bacteriol.* 40, 287–291. doi: 10.1099/00207713-40-3-287

van Der Palen, C. J., Slotboom, D. J., Jongejan, L., Reijnders, W. N., Harms, N., Duine, J. A., et al. (1995). Mutational analysis of mau genes involved in methylamine metabolism in *Paracoccus denitrificans*. *Eur. J. Biochem.* 230, 860–871. doi: 10.1111/j.1432-1033.1995.tb20629.x

Van Spanning, R. J., Wansell, C. W., De Boer, T., Hazelaar, M. J., Anazawa, H., Harms, N., et al. (1991). Isolation and characterization of the *moxJ*, *moxG*, *moxI*, and *moxR* genes of *Paracoccus denitrificans*: inactivation of *moxJ*, *moxG*, and *moxR* and the resultant effect on methylotrophic growth. *J. Bacteriol.* 173, 6948–6961.

Wilson, S. M., Gleisten, M. P., and Donohue, T. J. (2008). Identification of proteins involved in formaldehyde metabolism by *Rhodobacter sphaeroides*. *Microbiology* 154, 296–305. doi: 10.1099/mic.0.2007/011346-0

Wood, A. P., and Kelly, D. P. (1977). Heterotrophic growth of *Thiobacillus* A2 on sugars and organic acids. *Arch. Microbiol.* 113, 257–264. doi: 10.1007/BF00492033

Zhu, Y., Jameson, E., Parslow, R. A., Lidbury, I., Fu, T., Dafforn, T. R., et al. (2014). Identification and characterization of trimethylamine N-oxide (TMAO) demethylase and TMAO permease in *Methylocella silvestris* BL2. *Environ. Microbiol.* 16, 3318–3330. doi: 10.1111/1462-2920.12585

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

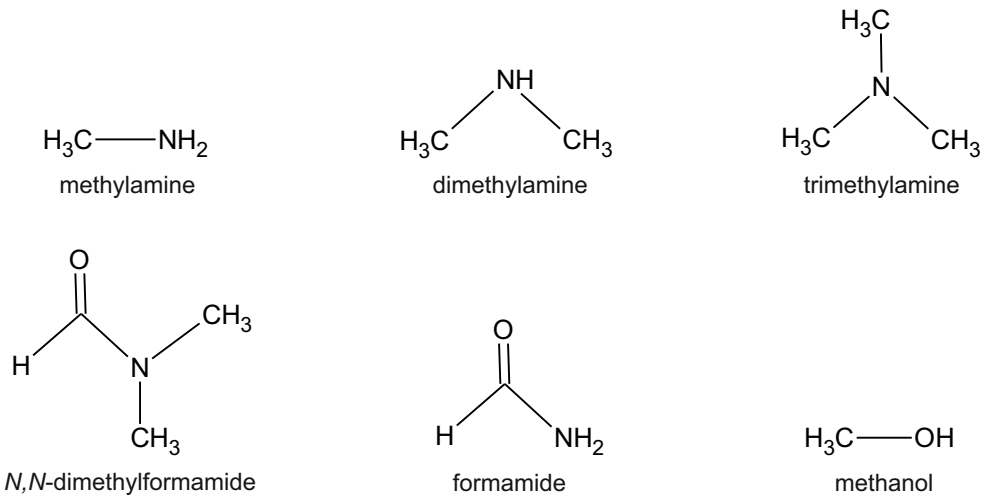


Figure S1 | Chemical structures of C1 compounds utilized by *P. aminophilus* JCM 7686.

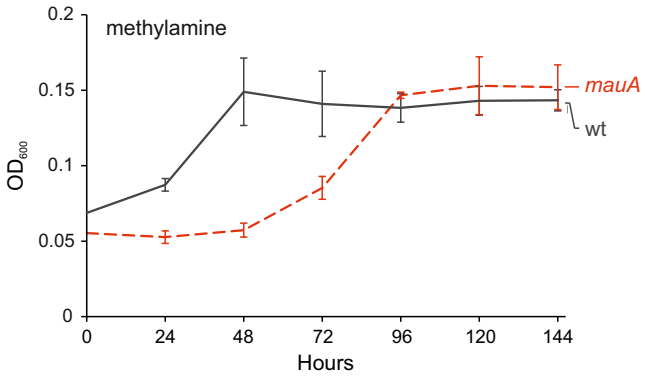


Figure S2 | Effect of the *mauA* mutation on growth of *P.aminophilus* JCM 7686 on methylamine. wt,wild type; *mauA*, *mauA* insertional mutant.The values are means of three replicates,and the error bars indicate the standard deviations.

Table S1.
Bacterial strains used in this study.

Strain	Characteristics	References
Escherichia coli TG1	F' [*traD36 proAB*$^{+}$ *lacI*^{q} *lacZΔM15*] *supE44 thi-1 Δ(lac-proAB) Δ(mcrB-hsdSM)5*; strain used for blue/white cloning and as a donor of plasmids in triparental mating	(Gibson, 1984)
E. coli S17-1	F^{-}, RP4-2(Km::Tn*7*,Tc::Mu-1), *pro*-82, *recA*1, *endA*1, *thiE*1, *hsdR*17, *creC*510, λ*pir* lysogen; strain used for maintaining pDS132 and its derivatives and as a donor of plasmids in biparental mating	(Simon et al., 1983)
Paracoccus aminophilus JCM 7686R	Rif^{r} derivative of wild-type strain JCM 7686	(Bartosik et al., 2002)
P. aminophilus xoxF	JCM 7686R derivative, *xoxF* (JCM7685_0090)::Km	This study
P. aminophilus xoxF compl.	JCM 7686R derivative, *xoxF* (JCM7685_0090)::Km, pBBRKm-*xoxF*	This study
P. aminophilus dmmA	JCM 7686R derivative, *dmmA* (JCM7685_ pAMI6p073)::Km	This study
P. aminophilus dmmB	JCM 7686R derivative, *dmmB* (JCM7685_ pAMI6p072)::Km	This study
P. aminophilus dmmC	JCM 7686R derivative, *dmmC* (JCM7685_ pAMI6p071)::Km	This study
P. aminophilus dmmD	JCM 7686R derivative, *dmmD* (JCM7685_ pAMI6p074)::Km	This study
P. aminophilus mauA	JCM 7686R derivative, *mauA* (JCM7685_ JCM7686_0163)::Km	This study
P. aminophilus tmm1	JCM 7686R derivative, *tmm1* (JCM7685_ pAMI6p076)::Km	This study
P. aminophilus tmm2	JCM 7686R derivative, *tmm2* (JCM7685_ pAMI6p102)::Km	This study
P. aminophilus tmm12	JCM 7686R derivative, *tmm1* (JCM7685_ pAMI6p076)::Tc, *tmm2* (JCM7685_ pAMI6p102)::Km	This study

References:

Bartosik, D., Baj, J., Bartosik, A.A., and Wlodarczyk, M. (2002). Characterization of the replicator region of megaplasmid pTAV3 of Paracoccus versutus and search for plasmid-encoded traits. Microbiology 148, 871-881.

Gibson, T.J. (1984). Studies on Epstein-Barr genome. PhD thesis, University of Cambridge.

Simon, R., Priefer, U., and Pühler, A. (1983). A broad host range mobilization system for invivo genetic-engineering - transposon mutagenesis in gram-negative bacteria. Bio-Technology 1, 784-791.

Table S2.
Plasmids used in this study.

Plasmid	Characteristics/construction details	Applications	Reference
pBBR1MCS-3	Tc^r, *ori* pBBR1 (broad host range, replication in both *Gamma-* and *Alphaproteobacteria*), *oriT* RK2 (mobilizable), *lacZ*α, MCS	blue/white cloning, complemention of mutations, delivered into *P. aminophilus* JCM 7686 cells via triparental mating	Kovacs et al., 1994
pDIY-KM	Km^r, Ap^r, *ori* ColE1	source of kanamycin resistance gene in construction of cassettes for mutagenesis	Dziewit et al., 2011
pDS132	Cm^r; *ori* R6K (narrow host range, replication only in *E. coli* λpir); *oriT* RK2 (mobilizable); *sacB* (a counter-selection gene enabling selection of double-cross mutants on appropriate medium with sucrose)	delivery of cassettes for mutagenesis into *P. aminophilus* JCM 7686R cells via biparental mating	Philippe et al., 2004
pKRP12	Tc^r, Ap^r, *ori* ColE1	source of tetracycline resistance gene in construction of cassettes for mutagenesis	Reece and Phillips, 1995
pRK2013	Km^r, the conjugal transfer system of RK2	helper plasmid used to mobilize mobilizable vectors in triparental mating	Ditta et al. 1980
pF4.1.D	pCC1FOS (Epicentre) derivative carrying 36-kb fragment of pAMI6 (fragment between 72,976 and 109,240 bp); the phosmid selected from phosmid genomic library of *P. aminophilus* JCM 7686 using PCR with LLPHMOT1 and LRPHMOT1 primers (Table S3)	source of *dmmDABC* operon for cloning	This study
pBBR-*dmmDABC*	pBBR1MCS-3 derivative carrying 4.8-kb restriction fragment of pF4.1.D (cut with StuI) including *dmmDABC*, the fragment was cloned into SmaI site of pBBR1MCS-3	source of *dmmABC* genes in construction of cassettes for mutagenesis	This study
pBBR-*xoxF*	pBBR1MCS-3 derivative carrying wild type *xoxF* gene - 2.0-kb PCR product obtained with LxoxKpn and RxoxXba primers (Table S3) cloned between KpnI and XbaI sites of pBBR1MCS-3)	complementation of *xoxF* mutation	This study
pDS-*dmmA*::Km	pDS132 derivative carrying 2.6-kb cassette for *dmmA* mutagenesis (Km^r); the cassette was constructed by cloning 1.9-kb restriction fragment of pBBR-*dmmDABC* (cut with PstI and SacI) containing *dmmA* gene into pDS132 (between PstI and SacI sites), resultant plasmid was then cut with NdeI and PvuI (creating 280-bp deletion in *dmmA* gene), 3'-sticky ends were blunted with T4 DNA polymerase and kanamycin resistance gene, cut from pDYI-KM with SmaI, was cloned into truncated *dmmA*	mutagenesis of *dmmA* gene	This study

Plasmid	Characteristics/construction details	Applications	Reference
pDS-*dmmB*::Km	pDS132 derivative carrying 4.0-kb cassette for *dmmB* mutagenesis (Km^r); the cassette was constructed by cloning 3.1-kb restriction fragment of pBBR-*dmmDABC* (cut with PstI) containing *dmmB* gene into pDS132 (into PstI site), resultant plasmid was then cut with ApaLI and RsrII (creating 400-bp deletion in *dmmB* gene), 5'-sticky ends were blunted with DNA polymerase I Klenow fragment and kanamycin resistance gene, cut from pDIY-KM with SmaI, was cloned into truncated *dmmB* gene	mutagenesis of *dmmB* gene	This study
pDS-*dmmC*::Km	pDS132 derivative carrying 2.0-kb DNA cassette for *dmmC* mutagenesis (Km^r); the cassette was constructed by cloning 1.7-kb restriction fragment of pBBR-*dmmDABC* (cut with PstI and SacI) into pDS132 (between PstI and SacI sites), resultant plasmid was then cut with PvuI and NdeI (creating 600-bp deletion in *dmmC* gene), 3'-sticky ends were blunted with T4 DNA polymerase and kanamycin resistance gene, cut from pDIY-KM with SmaI, was cloned into truncated *dmmC* gene	mutagenesis of *dmmC* gene	This study
pDS-*dmmD*::Km	pDS132 derivative carrying 1.8-kb cassette for *dmmD* mutagenesis (Km^r) constructed by overlap extension PCR with appropriate primers and cloned between two XbaI sites of the vector	mutagenesis of *dmmD* gene	This study
pDS-*mauA*::Km	pDS132 derivative carrying 1.7-kb cassette for *mauA* mutagenesis (Km^r) constructed by overlap extension PCR with appropriate primers (Table S3) and cloned between XbaI and SacI sites of the vector	mutagenesis od *mauA* gene	This study
pDS-*tmm1*::Km	pDS132 derivative carrying 1.8-kb cassette for *tmm1* mutagenesis (Km^r) constructed by overlap extension PCR with appropriate primers (Table S3) and cloned between XbaI and SacI sites of the vector	mutagenesis of *tmm1* gene	This study
pDS-*tmm1*::Tc	pDS132 derivative carrying 2.5-kb cassette for *tmm1* mutagenesis (Tc^r) constructed by overlap extension PCR with appropriate primers (Table S3) and cloned between XbaI and SacI sites of the vector	mutagenesis of *tmm1* gene	This study
pDS-*tmm2*::Km	pDS132 derivative carrying 1.8-kb cassette for *tmm2* mutagenesis (Km^r) constructed by overlap extension PCR with appropriate primers (Table S3) and cloned between XbaI and SacI sites of the vector	mutagenesis of *tmm2* gene	This study
pDS-*xoxF*::Km	pDS132 derivative carrying 2.4-kb cassette for *xoxF* mutagenesis (Km^r); the cassette was constructed by cloning 2.1-kb SacI restriction fragment of PCR product obtainted with 1xox and 6xox primers (Table S3) into pDS132 (into SacI site),	mutagenesis of *xoxF* gene	This study

Plasmid	Characteristics/construction details	Applications	Reference
	resultant plasmid was then cut with PvuI (creating 620-bp deletion in *xoxF* gene), 3'-sticky ends were blunted with T4 DNA polymerase and kanamycin resistance gene, cut from pDYI-KM with SmaI, was cloned into truncated *xoxF* gene		

Table S3.
Oligonucleotide primers used in this study.

Name	Sequence	Application
1xox	ATGAGCTCAAATCGGTCTCGGCCTTCGC	construction of cassette for mutagenesis of *xoxF*
6xox	ATGAGCTCCGGCCTTGCTGACATCTTGC	
0xox	CGTGGAGCCCGCATTTCAAG	confirmation of correct introduction of *xoxF* mutation
7xox	GGGCGAGCAGATAAGCGTAG	
LxoxKpn	ATGGTACCCCGGATGCAGGAAGGTTACG	amplification of wild type *xoxF* gene for complementation
RxoxXba	GCTCTAGAGGCCTAACTGCCTCTTAGCC	
LLPHMOT1	CGTCTAGAATGACCAAACGAGTGG	construction of cassette for mutagenesis of *tmm1* (Km^r) using overlap extension PCR
LRPHMOT1	TGAGACACAACGTGGCATCATAGG	
KLPHMOT1	CCTATGATGCCACGTTGTGTCTCA	
KRPHMOT1	ACTGATCCTGCGCCAGTGTTACAA	
RLPHMOT1	TTGTAACACTGGCGCAGGATCAGT	
RRPHMOT1	TAGAGCTCAGTTGCGCAGGTAG	
LRMOT1TC	GACACTATAGAACCATGCGGATCG	construction of cassette for mutagenesis of *tmm1* (Tc^r) using overlap extension PCR (LLPHMOT1 and RRPHMOT1 were also used for this purpose)
TLPHMOT1	CGATCCGCATGGTTCTATAGTGTC	
TRPHMOT1	TGAACCACTGATGTTGGTTTGCGC	
RLMOT1TC	GCGCAAACCAACATCAGTGGTTCA	
LMOT1XB	TAGGTACCGCGTGACGGCAAACAGGTTC	confirmation of correct introduction of *tmm1* mutations and amplification of wild type *tmm1* gene for complementation
RMOT1KP	GATCTAGAGGATCAGTTGCGCAGGTAGC	
LLPHMOT2	CGTCTAGAATGACGAAGCGAGTCGA	construction of cassette for mutagenesis of *tmm2*
LRPHMOT2	TGAGACACAACGGCGGATCGAACT	
KLPHMOT2	AGTTCGATCCGCCGTTGTGTCTCA	
KRPHMOT2	CACTGATCCTGCGCCAGTGTTACA	
RLPHMOT2	TGTAACACTGGCGCAGGATCAGTG	
RRPHMOT2	GCGAGCTCCAGTTCTTCAGATAGG	
LMOT2KP	TAGGTACCCAGATTTCGCGGCAGGGTTC	confirmation of correct introduction of *tmm2* mutation and amplification of wild type *tmm2* gene for complementation
RMOT2XB	GATCTAGACCGCAGAGTCACGTCAGTTC	

Name	Sequence	Application
1dmmami6	GATCTAGACTTTGGACGACGCCGATCAG	construction of cassette for mutagenesis of *dmmD*
2dmmami6	TTTGAGACACAACGTCCGCAGCGCCAAG	
3dmmami6	CTTGGCGCTGCGGACGTTGTGTCTCAAA	
4dmmami6	GCTCGAGGACTGGAGCCAGTGTTACAAC	
5dmmami6	GTTGTAACACTGGCTCCAGTCCTCGAGC	
6dmmami6	AGTCTAGAGCCTATCGCCATTGGGTGAC	
0dmmami6	CGGTGATGCCTTTGCAATGG	confirmation of correct introduction of *dmmD* mutation
7dmmami6	TTTCGTCAATGCGCTGTTCC	
LLPHMAUA	CGTCTAGAACCGAATCAGGCCTGTT	construction of cassette for mutagenesis of *mauA* using overlap extension PCR
LRPHMAUA	TGAGACACAACGAAAGCCAGCTGA	
KLPHMAUA	TCAGCTGGCTTTCGTTGTGTCTCA	
KRPHMAUA	GGAGGATAATCGGCCAGTGTTACA	
RLPHMAUA	GTAACACTGGCCGATTATCCTCCC	
RRPHMAUA	TCGAGCTCGATCATCAAGTCGATC	
LMAUA	ATCGACGCAGGGTCGCTTTC	confirmation of correct introduction of *mauA* mutation
RMAUA	TCGAGCTGCCCGATTACTTC	
Tmm13L	GTGCCTGTGCCGCTTTCTTC	RT-qPCR for transcript of *tmm1* gene
Tmm13R	GTCTTCCAGCGCTTCGTTCC	
Tmm23L	AGGGCGTCGTCTACGTCAAC	RT-qPCR for transcript of *tmm2* gene
Tmm23R	TCCGGCCCAGGATAATGTCG	
AceA1L	CGAAACCGACACGCCGAATG	RT-qPCR for transcript of *aceA* gene
AceA1R	GCACCTGTTTGCGCAGGTTG	
Hpr2L	CGGCGAGCCGAACATCAATC	RT-qPCR for transcript of *hpr* gene
Hpr2R	TCACGGCCGGCAAAGAAGTC	
Gck3L	CGCGCTGCTGATTTCCGATG	RT-qPCR for transcript of *gck* gene
Gck3R	CGTCGAGATGCCAGCGTTTG	
Ecm3L	CGCACCTATGCGGGTCATTC	RT-qPCR for transcript of *ecm* gene
Ecm3R	GTCATAGCCGGTCTGGGTTG	

Name	Sequence	Application
Mcm2L	ACGCCTTCTATCGCCGCAAC	RT-qPCR for transcript of *mcm* gene
Mcm2R	TGCCGACATCGCCTTCAACG	
RpoA3L	TCAAGGGCGTGACGCTGAAG	RT-qPCR for transcript of *rpoA* gene (reference)
RpoA3R	GTGATCGCGGTTCAGGATGG	

Table S4.
Genes of *Paracoccus aminophilus* JMC 7686 involved in the metabolism of C1 compounds.

Gene(s)	Replicon(s)	Encoded enzyme	EC number
		Oxidation of C1 substrates	
JCM7686_pAMI6p076 JCM7686_pAMI6p102	pAMI6	trimethylamine monooxygenase (Tmm)	1.14.13.8
JCM7686_pAMI6p069	pAMI6	trimethylamine-*N*-oxide demethylase (Tdm)	4.1.2.32
JCM7686_pAMI2p015-016	pAMI2	*N,N*-dimethylformamidase (DmfA1A2)	3.5.1.56
JCM7686_pAMI6p071-074	pAMI6	putative dimethylamine monooxygenase (DmmABCD)[1]	-
JCM7686_0162-0171	chromosome	methylamine dehydrogenase (MauFAEDBCJGMN)	1.4.9.1
JCM7686_pAMI6p105	pAMI6	glutamate-methylamine ligase (GmaS)	6.3.4.12
JCM7686_pAMI6p106-108	pAMI6	*N*-methyl-L-glutamate synthase (MgsABC)	2.1.1.21
JCM7686_pAMI6p096-099	pAMI6	*N*-methyl-L-glutamate dehydrogenase (MgdABCD)	1.5.99.5
JCM7686_0090	chromosome	methanol dehydrogenase (XoxF)	1.1.2.7
JCM7686_0085	chromosome	*S*-(hydroxymethyl)glutathione synthase (Gfa)	4.4.1.22
JCM7686_0086	chromosome	*S*-(hydroxymethyl)glutathione dehydrogenase (FlhA)	1.1.1.284
JCM7686_0089	chromosome	*S*-formylglutathione hydrolase (FghA)	3.1.2.12
JCM7686_pAMI6p077 JCM7686_pAMI6p095 JCM7686_pAMI5p256	pAMI6, pAMI5	5,10-methylene-tetrahydrofolate dehydrogenase/ 5,10-methylene-tetrahydrofolate cyclohydrolase (FolD)	1.5.1.5/ 3.5.4.9
JCM7686_pAMI6p078 JCM7686_pAMI6p094 JCM7686_pAMI5p257	pAMI6, pAMI5	formyltetrahydrofolate deformylase (PurU)	3.5.1.10

Gene(s)	Replicon(s)	Encoded enzyme	EC number
		Oxidation of C1 substrates	
JCM7686_pAMI6p042	pAMI6	formate-tetrahydrofolate ligase (FtfL)	6.3.4.3
JCM7686_pAMI4p036 JCM7686_1450	chromosome, pAMI4	formamidase (FmdA)	3.5.1.49
JCM7686_0639-0643 JCM7686_3476-3480 JCM7686_2088 JCM7686_pAMI1p027	chromosome, pAMI1	formate dehydrogenase (Fdh)	1.2.1.2
		Serine cycle	
JCM7686_2770 JCM7686_2900	chromosome	glycine hydroxymethyltransferase (GlyA)	2.1.2.1
JCM7686_1647 JCM7686_2765	chromosome	serine-glyoxylate transaminase (Sga)	2.6.1.45
JCM7686_2768	chromosome	hydroxypyruvate reductase (Hpr)	1.1.1.29
JCM7686_pAMI5p028	pAMI5	glycerate 2-kinase (Gck)	2.7.1.165
JCM7686_1636	chromosome	enolase (Eno)	4.2.1.11
JCM7686_2769	chromosome	phosphoenolpyruvate carboxylase (Ppc)	4.1.1.31
JCM7686_2576	chromosome	malate dehydrogenase (Mdh)	1.1.1.37
JCM7686_2766-2767	chromosome	malate-CoA ligase (MtkAB)	6.2.1.9
JCM7686_2763 JCM7686_3168	chromosome	malyl-CoA lyase (Mcl)[2]	4.1.3.24
		Ethylmalonyl-CoA pathway	
JCM7686_0515	chromosome	acetyl-CoA C-acetyltransferase (PhaA)	2.3.1.9

Gene(s)	Replicon(s)	Encoded enzyme	EC number
		Ethylmalonyl-CoA pathway	
JCM7686_0514	chromosome	acetoacetyl-CoA reductase (PhaB)	1.1.1.36
JCM7686_1380	chromosome	(*R*)-3-hydroxybutyryl-CoA dehydratase (CroR)	4.2.1.55
JCM7686_1036	chromosome	crotonyl-CoA carboxylase/reductase (Ccr)	1.3.1.85
JCM7686_2438	chromosome	ethylmalonyl-CoA epimerase (Epi)[3]	5.1.99.1
JCM7686_1037	chromosome	ethylmalonyl-CoA mutase (Ecm)	-
JCM7686_0656	chromosome	(2*S*)-methylsuccinyl-CoA dehydrogenase (Msd)	-
JCM7686_2571	chromosome	mesaconyl-CoA hydratase (Mcd)	4.2.1.148
JCM7686_2763 JCM7686_3168	chromosome	β-methylmalyl-CoA lyase (Mcl)[2]	4.1.3.24
JCM7686_1909 JCM7686_1905	chromosome	propionyl-CoA carboxylase (PccAB)	6.4.1.3
JCM7686_2438	chromosome	methylmalonyl-CoA epimerase (Epi)[3]	5.1.99.1
JCM7686_1912	chromosome	methylmalonyl-CoA mutase (Mcm)	5.4.99.2
		Glyoxylate shunt	
JCM7686_pAMI6p121	pAMI6	isocitrate lyase (AceA)	4.1.3.1
JCM7686_pAMI6p120	pAMI6	malate synthase (AceB)[4]	2.3.3.9
JCM7686_1627	chromosome	malate synthase G (GlcB)[4]	2.3.3.9

[1] Dimethylamine monooxygase activity of the purified products of the genes was not demonstrated. The role of the genes and their close homologs from *Methylocella silvestris* in dimethylamine metabolism was confirmed only by growth analyses of knockout mutants conducted in this study and by Zhu and co-workers (Zhu et al., 2014).
[2] Mcl is an enzyme catalyzing both the cleavage of malyl-CoA into glyoxylate and acetyl-CoA and the cleavage of β-methylmalyl-CoA into glyoxylate and propionyl-CoA (Erb et al., 2010).
[3] Epi is an enzyme catalyzing the conversion of both (2S)-ethylmalonyl-CoA and (2S)-methylmalonyl-CoA into their 2R-stereoisomers (Erb et al., 2008).
[4] The genome of *Paracoccus aminophilus* JCM 7686 encodes two putative non-homologous malate synthases.

Oświadczenia współautorów

Publikacja 1

Czarnecki J, Dziewit L, Kowalski L, Ochnio M, Bartosik D. 2015. Maintenance and genetic load of plasmid pKON1 of *Paracoccus kondratievae*, containing a highly efficient toxin-antitoxin module of the *hipAB* family. Plasmid 80:45-53.

Jakub Czarnecki
Zakład Genetyki Bakterii, Instytut Mikrobiologii,
Wydział Biologii, Uniwersytet Warszawski

Warszawa, 08.09.2015

OŚWIADCZENIE

Oświadczam, że w pracy:

Czarnecki J., L. Dziewit, L. Kowalski, M. Ochnio, D. Bartosik. 2015. Maintenance and genetic load pKON1 of Paracoccus kondratievae, *containing a highly efficient toxin-antitoxin module of the* hipAB *family. Plasmid 80: 45-53.*

mój udział polegał na (i) zaplanowaniu i nadzorowaniu prowadzonych prac eksperymentalnych (wraz z prof. dr. hab. Dariuszem Bartosikiem), (ii) analizie bioinformatycznej sekwencji plazmidu pKON1 (szczegółowa adnotacja sekwencji, identyfikacja elementów transpozycyjnych, analizy porównawcze i filogenetyczne modułów genetycznych pKON1, przygotowanie sekwencji nukleotydowej pKON1 do zdeponowania w bazie GenBank), (iii) identyfikacji *locus* stabilizującego pKON1, (iv) zbadaniu rozpowszechnienia systemów toksyna-antykoksyna (TA) z rodziny *hipAB* w genomach bakterii (baza NCBI), (v) sklonowaniu systemów TA, pochodzących z plazmidów pKON1 i pAMI8 oraz z chromosomu *Paracoccus aminophilus* JCM 7686, w wektorze wahadłowym pABW3, (vi) zbadaniu efektu stabilizującego poszczególnych systemów z rodziny *hipAB* w czterech szczepach z rodzaju *Paracoccus* (wraz z Łukaszem Kowalskim), (vii) analizie uzyskanych danych oraz przygotowaniu rycin i dodatkowych materiałów załączonych do pracy (*on-line supplementary materials*), a także napisaniu manuskryptu (wraz z prof. dr. hab. D. Bartosikiem). Badania te realizowałem w ramach pracy doktorskiej wykonywanej w Zakładzie Genetyki Bakterii Wydziału Biologii UW.

Podpis

.....

dr Łukasz Dziewit
Zakład Genetyki Bakterii, Instytut Mikrobiologii,
Wydział Biologii, Uniwersytet Warszawski

Warszawa, 08.09.2015

OŚWIADCZENIE

Oświadczam, że w pracy:

Czarnecki J., L. Dziewit, L. Kowalski, M. Ochnio, D. Bartosik. 2015. Maintenance and genetic load pKON1 of Paracoccus kondratievae, *containing a highly efficient toxin-antitoxin module of the* hipAB *family. Plasmid 80: 45-53.*

mój udział polegał na przygotowaniu DNA plazmidu pKON1 do sekwencjonowania, przeprowadzeniu pierwszej adnotacji sekwencji nukleotydowej pKON1 oraz wstępnych analiz porównawczych otrzymanej sekwencji, a także zdeponowaniu sekwencji w bazie GenBank.

Podpis

.....

Łukasz Kowalski
Zakład Genetyki Bakterii, Instytut Mikrobiologii,
Wydział Biologii, Uniwersytet Warszawski

Warszawa, 08.09.2015

OŚWIADCZENIE

Oświadczam, że w pracy:

Czarnecki J., L. Dziewit, L. Kowalski, M. Ochnio, D. Bartosik. 2015. Maintenance and genetic load pKON1 of Paracoccus kondratievae, *containing a highly efficient toxin-antitoxin module of the* hipAB *family. Plasmid 80: 45-53.*

uczestniczyłem w badaniu stabilności plazmidów niosących systemy toksyna-antytoksyna z rodziny *hipAB*, w różnych bakteriach z rodzaju *Paracoccus*.

Podpis

.....

Magdalena Ochnio
Zakład Genetyki Bakterii, Instytut Mikrobiologii,
Wydział Biologii, Uniwersytet Warszawski

Warszawa, 08.09.2015

OŚWIADCZENIE

Oświadczam, że w pracy:

Czarnecki J., L. Dziewit, L. Kowalski, M. Ochnio, D. Bartosik. 2015. Maintenance and genetic load pKON1 of Paracoccus kondratievae, *containing a highly efficient toxin-antitoxin module of the* hipAB *family. Plasmid 80: 45-53.*

mój udział polegał na konstrukcji wektora wahadłowego pJCB1, zawierającego system replikacyjny i system partycyjny plazmidu pKON1.

Podpis

.....

Prof. dr hab. Dariusz Bartosik
Zakład Genetyki Bakterii, Instytut Mikrobiologii,
Wydział Biologii, Uniwersytet Warszawski

Warszawa, 08.09.2015

OŚWIADCZENIE

Oświadczam, że w pracy:

Czarnecki J., L. Dziewit, L. Kowalski, M. Ochnio, D. Bartosik. 2015. Maintenance and genetic load pKON1 of Paracoccus kondratievae, *containing a highly efficient toxin-antitoxin module of the* hipAB *family. Plasmid 80: 45-53.*

mój udział polegał na zaplanowaniu i nadzorowaniu prowadzonych prac eksperymentalnych (wraz z Jakubem Czarneckim), analizie uzyskanych danych oraz przygotowaniu manuskryptu publikacji (wraz z J. Czarneckim).

Podpis
.....

Publikacja 2

Dziewit L, Czarnecki J, Wibberg D, Radlinska M, Mrozek P, Szymczak M, Schlüter A, Pühler A, Bartosik D. 2014. Architecture and functions of a multipartite genome of the methylotrophic bacterium *Paracoccus aminophilus* JCM 7686, containing primary and secondary chromids. BMC Genomics 15:124.

dr Łukasz Dziewit
Zakład Genetyki Bakterii, Instytut Mikrobiologii,
Wydział Biologii, Uniwersytet Warszawski

Warszawa, 08.09.2015

OŚWIADCZENIE

Oświadczam, że w pracy:

Dziewit L., J. Czarnecki, D. Wibberg, M. Radlinska, P. Mrozek, M. Szymczak, A. Schlüter, A. Pühler and D. Bartosik. Architecture and functions of a multipartite genome of the methylotrophic bacterium Paracoccus aminophilus *JCM 7686, containing primary and secondary chromids. BMC Genomics 15: 124.*

mój udział polegał na współautorstwie koncepcji badań, zaplanowaniu doświadczeń, wykonaniu części doświadczeń, sekwencjonowaniu DNA, adnotacji genomu, analizie i zdeponowaniu w bazie GenBank (NCBI) sekwencji genomu szczepu JCM 7686, analizie i interpretacji wyników badań, udziale w przygotowaniu manuskryptu (przygotowanie jego pierwszej wersji), pozyskaniu finansowania badań (granty nr IP2010 008670 i IP2011 011471); przygotowaniu odpowiedzi na uwagi recenzentów.

Podpis

Jakub Czarnecki
Zakład Genetyki Bakterii, Instytut Mikrobiologii,
Wydział Biologii, Uniwersytet Warszawski

Warszawa, 08.09.2015

OŚWIADCZENIE

Oświadczam, że w pracy:

Dziewit L., J. Czarnecki, D. Wibberg, M. Radlinska, P. Mrozek, M. Szymczak, A. Schlüter, A. Pühler and D. Bartosik. Architecture and functions of a multipartite genome of the methylotrophic bacterium Paracoccus aminophilus *JCM 7686, containing primary and secondary chromids. BMC Genomics 15: 124.*

mój udział polegał na analizie strukturalnej oraz funkcjonalnej plazmidów i chromidów *Paracoccus aminophilus* JCM 7686.

(i) konstrukcji mapy fizycznej chromosomu *P. aminophilus* JCM 7686 (wraz z dr. Łukaszem Dziewitem), (ii) konstrukcji mobilizowalnych wektorów wahadłowych, zawierających systemy replikacyjne replikonów pozachromosomowych *P. aminophilus*, (iii) konstrukcji szczepów bezplazmidowych i identyfikacji replikonów niezbędnych, (iv) analizie kinetyki wzrostu szczepów bezplazmidowych na podłożach pełnych i minimalnych, (v) zbadaniu funkcji genu *ccrM P. aminophilus* (analiza mutacyjna i komplementacja mutacji). Badania te realizowałem w ramach pracy doktorskiej wykonywanej w Zakładzie Genetyki Bakterii UW.

Podpis

.....

Daniel Wibberg
Center for Biotechnology,
Bielefeld University,
Bielefeld, Germany

Bielefeld, 10.03.2015

CO-AUTHOR STATEMENT

I hereby certify that in the publication:

Dziewit L., J. Czarnecki, D. Wibberg, M. Radlinska, P. Mrozek, M. Szymczak, A. Schlüter, A. Pühler and D. Bartosik. Architecture and functions of a multipartite genome of the methylotrophic bacterium Paracoccus aminophilus *JCM 7686, containing primary and secondary chromids. BMC Genomics 15: 124.*

my contribution comprised revising it critically for important intellectual content, drafting selected parts of the manuscript, as well as participating in genome manual annotation and its structural analyses.

Signature

Wibberg

Dr Monika Radlińska
Zakład Wirusologii, Instytut Mikrobiologii,
Wydział Biologii, Uniwersytet Warszawski

Warszawa, 08.09.2015

OŚWIADCZENIE

Oświadczam, że w pracy:

Dziewit L., J. Czarnecki, D. Wibberg, M. Radlinska, P. Mrozek, M. Szymczak, A. Schlüter, A. Pühler and D. Bartosik. Architecture and functions of a multipartite genome of the methylotrophic bacterium Paracoccus aminophilus *JCM 7686, containing primary and secondary chromids. BMC Genomics 15: 124.*

mój udział polegał na badaniu specyficzności metylotransferaz kodowanych w genomie *P. aminophilus* JCM 7686 oraz indukcji i identyfikacji funkcjonalnego faga ΦPam-6.

Podpis

.....

Paulina Mrozek
Uniwersytet Warszawski

Monachium, 14.02.2014

OŚWIADCZENIE

Oświadczam, że w pracy:

Dziewit L., J. Czarnecki, D. Wibberg, M. Radlinska, P. Mrozek, M. Szymczak, A. Schlüter, A. Pühler and D. Bartosik. Architecture and functions of a multipartite genome of the methylotrophic bacterium Paracoccus aminophilus *JCM 7686, containing primary and secondary chromids. BMC Genomics 15: 124.*

mój udział polegał na badaniach aktywności metylotransferaz kodowanych w genomie *Paracoccus aminophilus* JCM 7686. Badania te realizowałam w ramach pracy magisterskiej wykonywanej w Zakładzie Wirusologii UW.

Podpis

Paulina Mrozek

Michał Szymczak
Uniwersytet Warszawski

Warszawa, 18.02.2014

OŚWIADCZENIE

Oświadczam, że w pracy:

Dziewit L., J. Czarnecki, D. Wibberg, M. Radlinska, P. Mrozek, M. Szymczak, A. Schlüter, A. Pühler and D. Bartosik. Architecture and functions of a multipartite genome of the methylotrophic bacterium Paracoccus aminophilus *JCM 7686, containing primary and secondary chromids. BMC Genomics 15: 124.*

mój udział polegał na badaniu zakresu gospodarzy plazmidów *Paracoccus aminophilus* JCM 7686. Badania te realizowałem w ramach pracy magisterskiej wykonywanej w Zakładzie Genetyki Bakterii UW.

Podpis

Szymczak
.....

Dr. Andreas Schlüter
Center for Biotechnology,
Bielefeld University,
Bielefeld, Germany

Bielefeld, 10.03.2015

CO-AUTHOR STATEMENT

I hereby certify that in the publication:

Dziewit L., J. Czarnecki, D. Wibberg, M. Radlinska, P. Mrozek, M. Szymczak, A. Schlüter, A. Pühler and D. Bartosik. Architecture and functions of a multipartite genome of the methylotrophic bacterium Paracoccus aminophilus *JCM 7686, containing primary and secondary chromids. BMC Genomics 15: 124.*

my contribution comprised revising it critically for important intellectual content, providing the access to genomic analysis tools and supervising those analyses.

Signature

Prof. Dr. Alfred Pühler
Center for Biotechnology,
Bielefeld University,
Bielefeld, Germany

Bielefeld, 10.03.2015

CO-AUTHOR STATEMENT

I hereby certify that in the publication:

Dziewit L., J. Czarnecki, D. Wibberg, M. Radlinska, P. Mrozek, M. Szymczak, A. Schlüter, A. Pühler and D. Bartosik. Architecture and functions of a multipartite genome of the methylotrophic bacterium Paracoccus aminophilus *JCM 7686, containing primary and secondary chromids. BMC Genomics 15: 124.*

my contribution comprised revising it critically for important intellectual content, providing the access to genomic analysis tools and supervising those analyses.

Signature

.....

Prof. Dr. A. Pühler

Prof. dr hab. Dariusz Bartosik
Zakład Genetyki Bakterii, Instytut Mikrobiologii,
Wydział Biologii, Uniwersytet Warszawski

Warszawa, 08.09.2015

OŚWIADCZENIE

Oświadczam, że w pracy:

Dziewit L., J. Czarnecki, D. Wibberg, M. Radlinska, P. Mrozek, M. Szymczak, A. Schlüter, A. Pühler and D. Bartosik. Architecture and functions of a multipartite genome of the methylotrophic bacterium Paracoccus aminophilus *JCM 7686, containing primary and secondary chromids. BMC Genomics 15: 124.*

mój udział polegał na nadzorowaniu prowadzonych prac eksperymentalnych, zaplanowaniu eksperymentów (wraz z dr. Łukaszem Dziewitem) oraz przygotowaniu manuskryptu publikacji (wraz z dr. Ł. Dziewitem i Danielem Wibbergiem).

Podpis

.....

Publikacja 3

Dziewit L, Czarnecki J, Prochwicz E, Wibberg D, Schlüter A, Pühler A, Bartosik D. 2015. Genome-guided insight into the methylotrophy of *Paracoccus aminophilus* JCM 7686. Front Microbiol 6:852.

dr Łukasz Dziewit
Zakład Genetyki Bakterii, Instytut Mikrobiologii,
Wydział Biologii, Uniwersytet Warszawski

Warszawa, 08.09.2015

OŚWIADCZENIE

Oświadczam, że w pracy:

Dziewit L., J. Czarnecki, E. Prochwicz, D. Wibberg, A. Schlüter, A. Pühler and D. Bartosik. 2015. Genome-guided insight into the methylotrophy of Paracoccus aminophilus *JCM 7686. Frontiers in Microbiology 6: 852.*

mój udział polegał na analizie bioinformatycznej sekwencji genomu *Paracoccus aminophilus* JCM 7686 pod kątem identyfikacji genów zaangażowanych w metylotrofię (wraz z Jakubem Czarneckim), zaplanowaniu eksperymentów (wraz z Jakubem Czarneckim), nadzorowaniu prowadzonych prac eksperymentalnych (wraz z prof. dr. hab. Dariuszem Bartosikiem i Jakubem Czarneckim), analizie rozpowszechnienia genów związanych z metylotrofią w bakteriach z rodzaju *Paracoccus* (wraz z Danielem Wibbergiem i Jakubem Czarneckim), analizie uzyskanych danych (wraz z Jakubem Czarnecki i prof. dr. hab. Dariuszem Bartosikiem) oraz przygotowaniu manuskryptu (wraz z Jakubem Czarneckim i prof. dr. hab. Dariuszem Bartosikiem).

Podpis

Dziewit

.....

Jakub Czarnecki
Zakład Genetyki Bakterii, Instytut Mikrobiologii,
Wydział Biologii, Uniwersytet Warszawski

Warszawa, 08.09.2015

OŚWIADCZENIE

Oświadczam, że w pracy:

Dziewit L., J. Czarnecki, E. Prochwicz, D. Wibberg, A. Schlüter, A. Pühler and D. Bartosik. 2015. Genome-guided insight into the methylotrophy of Paracoccus aminophilus *JCM 7686. Frontiers in Microbiology 6: 852.*

mój udział polegał na: (i) analizie bioinformatycznej sekwencji nukleotydowej genomu *Paracoccus aminophilus* JCM 7686, (ii) identyfikacji genów zaangażowanych w metylotrofię (wraz z dr. Łukaszem Dziewitem), (iii) zaplanowaniu eksperymentów (wraz z dr. Ł. Dziewitem), (iv) nadzorowaniu prac eksperymentalnych prowadzonych przez studentów zaangażowanych w realizację projektu, (v) zbadaniu kinetyki wzrostu *P. aminophilus* na różnych związkach C1, (vi) konstrukcji szczepów *P. aminophilus* zmutowanych w genach *xoxF*, *mauA*, *tmm1*, *tmm2*, *dmmA*, *dmmB*, *dmmC*, (vii) badaniu wpływu mutacji na wzrost uzyskanych szczepów na różnych związkach C1 (analizy wzrostowe i komplementacje mutacji), (vii) przeprowadzeniu analiz RT-qPCR wybranych genów *P. aminophilus* zaangażowanych w metabolizm metylotroficzny, (viii) analizie rozpowszechnienia genów związanych z metylotrofią w genomach bakterii z rodzaju *Paracoccus* (wraz z dr. Ł. Dziewitem i Danielem Wibbergiem), (ix) identyfikacji zespołu genów zaangażowanych w cykl serynowy i utlenianie metylowanych amin (w genomach *Paracoccus* sp. N5 i bakterii z kladu *Roseobacter*), (x) analizie uzyskanych danych i przygotowaniu manuskryptu (wraz z dr. Ł. Dziewitem i prof. dr. hab. D. Bartosikiem) oraz przygotowaniu rycin i dodatkowych materiałów załączonych do pracy (*on-line supplementary materials*) (wraz z dr. Ł. Dziewitem). Badania te realizowałem w ramach pracy doktorskiej wykonywanej w Zakładzie Genetyki Bakterii UW.

Podpis

.....

Emilia Prochwicz
Zakład Genetyki Bakterii, Instytut Mikrobiologii,
Wydział Biologii, Uniwersytet Warszawski

Warszawa, 08.09.2015

OŚWIADCZENIE

Oświadczam, że w pracy:

Dziewit L., J. Czarnecki, E. Prochwicz, D. Wibberg, A. Schlüter, A. Pühler and D. Bartosik. 2015. Genome-guided insight into the methylotrophy of Paracoccus aminophilus *JCM 7686. Frontiers in Microbiology 6: 852.*

mój udział polegał na przygotowaniu mutanta *Paracoccus aminophilus* JCM 7686 w genie *dmmD* oraz zbadaniu wpływu mutacji na wzrost tej bakterii na różnych związkach C1.

Podpis

Emilia Prochwicz

.....

Daniel Wibberg
Center for Biotechnology,
Bielefeld University,
Bielefeld, Germany

Bielefeld, 08.09.2015

CO-AUTHOR STATEMENT

I hereby certify that in the publication:

Dziewit L., J. Czarnecki, E. Prochwicz, D. Wibberg, A. Schlüter, A. Pühler and D. Bartosik. 2015. Genome-guided insight into the methylotrophy of Paracoccus aminophilus *JCM 7686. Frontiers in Microbiology 6: 852.*

my contribution comprised revising it critically for important intellectual content and participating in search for methylotrophy genes in genomes of *Paracoccus* spp.

Signature

.....

Dr. Andreas Schlüter
Center for Biotechnology,
Bielefeld University,
Bielefeld, Germany

Bielefeld, 08.09.2015

CO-AUTHOR STATEMENT

I hereby certify that in the publication:

Dziewit L., J. Czarnecki, E. Prochwicz, D. Wibberg, A. Schlüter, A. Pühler and D. Bartosik. 2015. Genome-guided insight into the methylotrophy of Paracoccus aminophilus *JCM 7686. Frontiers in Microbiology 6: 852.*

my contribution comprised revising it critically for important intellectual content, providing the access to genomic analysis tools and supervising those analyses.

Signature

.....

Prof. Dr. Alfred Pühler
Center for Biotechnology,
Bielefeld University,
Bielefeld, Germany

Bielefeld, 08.09.2015

CO-AUTHOR STATEMENT

I hereby certify that in the publication:

Dziewit L., J. Czarnecki, E. Prochwicz, D. Wibberg, A. Schlüter, A. Pühler and D. Bartosik. 2015. Genome-guided insight into the methylotrophy of Paracoccus aminophilus *JCM 7686. Frontiers in Microbiology 6: 852.*

my contribution comprised revising it critically for important intellectual content, providing the access to genomic analysis tools and supervising those analyses.

Signature

.....

Prof. dr hab. Dariusz Bartosik
Zakład Genetyki Bakterii, Instytut Mikrobiologii,
Wydział Biologii, Uniwersytet Warszawski

Warszawa, 08.09.2015

OŚWIADCZENIE

Oświadczam, że w pracy:

Dziewit L., J. Czarnecki, E. Prochwicz, D. Wibberg, A. Schlüter, A. Pühler and D. Bartosik. 2015. Genome-guided insight into the methylotrophy of Paracoccus aminophilus *JCM 7686. Frontiers in Microbiology 6: 852.*

mój udział polegał na nadzorowaniu prowadzonych prac eksperymentalnych (wraz z dr. Łukaszem Dziewitem i Jakubem Czarneckim), analizie uzyskanych wyników (wraz z Ł. Dziewitem i J. Czarneckim) oraz przygotowaniu manuskryptu (wraz z dr. Ł. Dziewitem i J. Czarneckim)

Podpis

.....